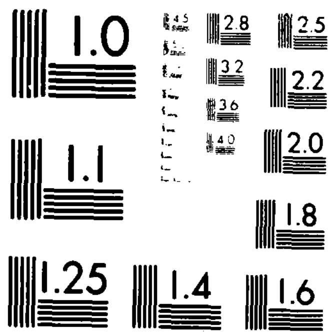
AD-A147 923 ESTIMATING WARTIME SUPPORT RESOURCE REQUIREMENTS STATISTICAL AND RELATED POLICY ISSUES(U) RAND CORP SANTA MONICA CA L B EMBRY JUL 84 RAND/N-2163-AF 1/2

UNCLASSIFIED F49620-82-C-0018 F/G 15/5 NL



MICROCOPY RESOLUTION TEST CHART

(12)

A RAND NOTE

ESTIMATING WARTIME SUPPORT RESOURCE REQUIREMENTS:
STATISTICAL AND RELATED POLICY ISSUES

Lloyd B. Embry

July 1984

N-2163-AF

Prepared for

The United States Air Force

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

The research reported here was sponsored by the Directorate of Operational Requirements, Deputy Chief of Staff/Research, Development, and Acquisition, Hq USAF, under Contract F49620-82-C-0018. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

Published by The Rand Corporation

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER N-2163-AF	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Estimating Wartime Support Resource Requirements: Statistical and Related Policy Issues		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Lloyd B. Embry		8. CONTRACT OR GRANT NUMBER(s) F49620-82-C-0018
9. PERFORMING ORGANIZATION NAME AND ADDRESS The Rand Corporation 1700 Main Street Santa Monica, CA. 90406		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Requirements, Programs & Studies Group (AF/RDQM) Ofc, DCS/R&D and Acquisition Hq USAF Washington, D.C. 20330		12. REPORT DATE July 1984
		13. NUMBER OF PAGES 94
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for Public Release: Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

No restrictions

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Spare parts
Failure
Logistics planning
Weapons systems
Military Aircraft
Aircraft engines
Military supplies

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

See reverse side

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

This Note addresses statistical and policy issues central to improving estimates of wartime support resource requirements. It uses the current problem of establishing the level of investment in spare engines for the C-5 aircraft to elucidate a number of these issues. The author examines the assumptions used to project peacetime experience to wartime activity levels and concludes that peacetime operational experiments, coupled with engineering projections of wartime failure rates, could be used to test these assumptions and provide an improved basis for resource requirements computations.

A RAND NOTE

ESTIMATING WARTIME SUPPORT RESOURCE REQUIREMENTS: STATISTICAL AND RELATED POLICY ISSUES

Lloyd B. Embry

July 1984

N-2163-AF

Prepared for The United States Air Force

Rand
1700 MAIN STREET
P.O. BOX 2138
SANTA MONICA, CA 90406-2138

PREFACE

The Air Force spends several billion dollars annually to procure the spare parts and other resources needed to support modern aircraft weapon systems. A large fraction of this investment is used to obtain assets for support of wartime operations. Because wartime data are not available to estimate the parameters used in computing wartime support requirements, these requirements are based on projections from peacetime experience.

The assumptions used to project peacetime experience to wartime activity levels have important resource implications, but little has been done to test them empirically. Peacetime operational experiments, coupled with engineering projections of wartime failure rates, could be used to test these assumptions and provide an improved basis for resource requirements computations.

This Note addresses statistical and policy issues central to improving estimates of wartime support resource requirements. It uses the current problem of establishing the level of investment in spare engines for the C-5 aircraft to elucidate a number of these issues. The results should be of interest to policymakers concerned with logistics resource allocation, operational commanders whose wartime capabilities are affected by statistical assumptions and related policy decisions, and the personnel responsible for producing requirements estimates.

This work was conducted as part of the study effort "The Driving Inputs and Assumptions of Stockage/Assessment Models" within the Project AIR FORCE Resource Management Program. The text of the Note has been accepted by the Rand Graduate Institute in partial fulfillment of the requirements for the doctoral degree in policy analysis.

SUMMARY

This Note addresses the issue of forecasting wartime demands for C-5 engines and proposes use of a combination of engineering projections and operational experience to improve demand estimates. The research was undertaken because the analytic basis for the key assumptions that are currently used to forecast wartime demand is very weak, even though these forecasts drive requirements for billions of dollars worth of support resources.

As a case in point, the Air Force budgeted over $3 billion in FY 1983 to buy the exchangeable[1] spare parts needed to support modern aircraft. Funding needs are projected to triple over the next few years. Although the high cost of spares has received a great deal of attention in the press lately, most of this publicity has dealt with only one cause of large spares budgets--the high unit costs charged by some government suppliers. The procedures used to determine procurement quantities represent another important explanation for high support costs.

The Air Force has completed a detailed study of the reasons for growth in spares requirements that focused primarily on methods used to develop requirements to support peacetime operations. Stocks are also purchased in peacetime to support the increased operating tempo expected in time of war. **War reserve needs represent a large fraction of the total "requirement" for any particular item. However, wartime data are not available to support computation of wartime requirements, so estimates are based on extrapolations from peacetime data.**

The assumptions used to project peacetime experience to wartime activity levels warrant careful scrutiny. Unfortunately, little effort has been devoted to this issue, and peacetime operating policies are not designed to generate the data that would be needed to test them empirically. The Air Force has reason to be concerned about the validity of some of these assumptions, and it should consider policy

---

[1]"Exchangeable" components can be repaired and restored to service after they fail.

changes that would make them less difficult to address. The contemporary problem of estimating the requirement for spare C-5 engines is used as a vehicle to address these issues.

STATISTICAL ASSUMPTIONS AND REQUIREMENTS ESTIMATES

The most critical assumption embedded in nearly all stockage requirements methodologies is that the demand for spare parts is proportional to the aircraft flying program. This relationship, which is often improperly referred to as the "linearity" assumption, drives estimates of the stock and maintenance resources needed to support the surge from peacetime to wartime levels of activity.

Over the years a considerable body of evidence has accumulated that refutes the "linearity" assumption. Several studies have found that the demand rate--the number of demands for spare parts recorded per flying hour--declines with increasing sortie lengths or aircraft utilization rates. This phenomenon has been observed in both military and civilian airline data.

The apparent inverse relationship between utilization and demand rates has been observed across a wide range of components, but it is particularly well documented for turbine engines used in transport aircraft. There is a straightforward engineering explanation for this phenomenon. Power transients, or engine cycles, contribute at least as much to the stress that leads to engine removal as do engine flying hours. Hence cycle density, or the cycles per flying hour ratio, should affect removal rates measured in removals per flying hour.

The training syllabus for fighter pilots is designed to approximate wartime sortie profiles. Nevertheless, cycle density for tactical aircraft might increase in wartime as pilots are forced to make frequent throttle changes in response to tactical developments. Average cycle density should decrease for bomber and cargo missions because these aircraft will not be making multiple landings during a single sortie as they do during many peacetime training missions. This decrease in cycle density implies a decrease in the wartime engine removal rate. If this effect is not offset by that of increased time at full throttle, which would tend to increase the removal rate, a linear extrapolation of peacetime demand experience could result in an overstatement of wartime resource requirements.

One obvious alternative to the linearity assumption is to develop wartime removal rates for different components based on an assessment of their dominant failure modes and analysis that links these modes to measures of aircraft utilization. This approach recognizes that:

- The component removal rate observed in peacetime reflects the combined result of a number of component failure modes,
- At least some of these failure modes can be associated with aircraft (and component) utilization,
- Utilization patterns in wartime will differ from peacetime experience,
- Peacetime simulation of wartime operating conditions can provide a basis for:
  Estimating the relationship between utilization parameters and component life,
  Developing logistics parameter estimates for use in wartime resource requirements computations.

SCHEDULING POLICIES AND LOGISTICS DATA

The C-5 is used for both cargo and training missions in peacetime. Thus there should be a distribution of average cycle lengths in the peacetime data that could shed light on the wartime requirements problem. However, day-to-day variability in utilization cannot be used to address this problem, because engine stress is cumulative. Variation in cycle density over engine "lifetimes"--measured from installation to removal--is needed to develop a relationship between engine removal rates and utilization.

The Air Force is planning to spend about $100 million for additional war reserve spare engines to support an expanded C-5 fleet. This resource requirement is driven by an estimate of the wartime demand rate. The original intent of this analysis was to exploit natural variability in peacetime mission demands to develop an improved wartime C-5 (TF-39) engine removal rate estimate.

Unfortunately, the scheduling rules used to assign aircraft to missions effectively eliminates variability in average cycle length over engine lifetimes, and the engine is undergoing a major modification program. **Currently it is not possible to estimate the wartime removal rate using peacetime data.**

Logistics data needs have not traditionally been considered in operational planning, but the resource implications of the wartime requirements forecasting problem suggest that this tradition should change. Commercial experience suggests that demand rates decrease with increasing cycle lengths. Despite the lack of variability in peacetime C-5 utilization, an actuarial ("life data") analysis of data describing normal peacetime operations produced results consistent with both commercial experience and engineering theory. Changing aircraft scheduling policies to generate the data needed for logistics parameter estimation would provide a basis for improved requirements forecasts.

The data needed to develop improved wartime demand estimates can most easily be developed from life tests of individual components. An experiment that could be conducted with only minor changes in aircraft scheduling policy would be similar to, but more extensive than, the qualification acceptance tests used for new military hardware.

The wartime demand estimation problem also provides an opportunity to combine use of engineering projections and operational experience. The engine manufacturer has developed an engineering model that could be used to estimate the wartime engine removal rate as a function of operational parameters. The model's predictions could be tested in an experiment in which different aircraft were assigned to training and cargo-hauling missions. In addition, Bayesian methods could be used to update the model's predictions during the course of the experiment.

The total cost of a program designed to improve wartime C-5 engine removal rate estimates would be considerably less than the cost of a single spare engine. The only incremental cost would be for applying the engineering model to the military problem; both training and cargo missions will be flown whether or not aircraft are assigned to generate data relevant to this critical problem. **The potential benefits**

of such a program could include savings in engine and engine spares procurement of over $100 million. Savings for other components, including those used on other types of aircraft, and for other support resources could increase this total considerably.

IMPROVING REQUIREMENTS ESTIMATES

Past studies suggest strongly that the linearity assumption overstates wartime demand for the components used on cargo aircraft, but it may understate demand for some components used in the tactical fleet. Nevertheless, the services continue to use this assumption in requirements computations for all types of aircraft.

There are undoubtedly many explanations for why research results have not prompted changes in computational methods. One of the most important is that the linearity assumption is considered conservative. A conservative assumption that overstates requirements for one type of resource, however, reduces the level of wartime capability that can be obtained for a limited budget; and defense budgets are always constrained, regardless of what the press would have us believe. Use of such conservative estimates can limit other aspects of combat capability through an overcommitment of limited resources to procurement of support resources for airlift.

Although the data currently available cannot be used to develop a wartime removal rate estimate for C-5 engines, the necessary data can be collected at fairly low cost. Increasing pressures on the Defense budget, compounded by internal pressures to restrain support cost growth, demand a rigorous examination of the assumptions that provide the basis for spares requirements estimates. **Hence the Air Force should implement a program to deal with the specific problem of C-5 engine requirements, as well as other programs to address the more general problem of estimating wartime resource requirements.**

ACKNOWLEDGMENTS

Michael Creasy's experience at Dover Air Force Base enabled him to address C-5 operations, data sources, and data quality issues with unusual clarity. Donald Johnson of the Oklahoma City Air Logistics Center extracted the data needed for the study. John Fitzgerald of Headquarters, Air Force Logistics Command explained the engine requirements computation and identified a body of past work that suggests that cycle density affects engine removal rates. Colonel William Smiley of the Air Staff served as project officer, and opened the door to several Air Force organizations.

James Duhig of the Lockheed-Georgia Company described Lockheed's engine maintenance analyses and provided a notional wartime route structure for the C-5 fleet. Richard Young, Manager of the TF39 Engine Program at General Electric, explained the results of the engine manufacturer's failure analyses. Patrick Sibley of the Northrop Corporation identified several cases in which cycle density appears to have affected the demand for engines in peacetime. Stanley Nowlan, former chief maintenance analyst for United Airlines, recommended an actuarial approach to the data analysis problem and summarized the results of some work done at United. Charles S. Smith of the System Development Foundation helped structure the discussion of the engine failure process, and Professor Richard Soland of George Washington University explained some of his past work in Bayesian estimation and renewal theory.

Rand colleagues William Rogers, Gordon Crawford, and John Abell provided a great deal of assistance with statistical problems. Michael Rich, director of the Project AIR FORCE Resource Management Program, contributed to the definition of the policy aspects of the requirements problem. Arthur Alexander's suggestion that the literature on engine monitoring systems should be relevant to the requirements problem led to the proposal to combine engineering estimates and operational data to generate wartime parameter estimates. Milton Kamins located several key references. Charles Wolf, Jr. arranged the funding needed to complete the study, and Carol Tripp, the "Dean for Students" at the Rand Graduate Institute, offered continuing emotional support.

I also wish to thank Gail Halverson for doing most of the initial data processing, and Dee Saenz and Kim Schlotterback for administrative support early in the study. Barbara Urwin patiently made many changes and prepared the final manuscript.

Any remaining errors of fact or interpretation are my responsibility.

CONTENTS

PREFACE iii

SUMMARY v

ACKNOWLEDGMENTS xi

FIGURES xv

TABLES xvii

Section

I. INTRODUCTION 1

- Statistical and Policy Issues That Affect Requirements Forecasts 2
- The C-5 Engine Requirements Problem 3
- Outline of the Report 5

II. ESTIMATING STOCKAGE REQUIREMENTS 6

- Weaknesses in the Assumptions Used To Develop Stockage Requirements 10
- Wartime Failure Rates of Aircraft Engines 12

III. AIRCRAFT TURBINE ENGINE RELIABILITY 13

- Reasons for Engine Removal 13
- Measuring Engine Stress 15
- Implications of the Mission Profile for the Engine Failure Rate 17
- Airline Engine Removal Experience 19
- An Engineering Model of the Engine Failure Process 21

IV. C-5 ENGINE UTILIZATION DATA AND EXPLORATORY DATA ANALYSIS . 24

V. ESTIMATING WARTIME REMOVAL RATES 32

- Statistical Analysis of Engine Utilization Data 33
- A Survival Analysis of Engine Removal Data 36
- Estimating the Parameters of a Weibull Distribution 42

VI. GENERATING DATA FOR LOGISTICS PARAMETER ESTIMATION 46

An Approach to the Experimental Design Problem 46
A "Clinical Trial" of Engine Treatment Alternatives 49
Testing and Updating Predictions of the GE OPSEV Model 56

VII. FORECASTING LOGISTICS RESOURCE REQUIREMENTS: POLICY IMPLICATIONS 59

Airlift Analogies to the C-5 Engine Requirements Problem .. 60
Analogies in Other Settings 62
Other Applications of Life Test Results 63
Engineering Projections vs. Operational Experience 64
Incentives for Change 65

Appendixes

A. SENSITIVITY OF REQUIREMENTS TO PARAMETER ESTIMATES 69

B. THE GE OPSEV MODEL 75

C. DATA PROCESSING 78

D. BAYESIAN ANALYSIS OF REMOVAL DATA 83

BIBLIOGRAPHY 89

FIGURES

1. The Aviation Logistics Support System 8

2. Requirements Sensitivity to Parameter Estimates 9

3. Typical Mission Sortie Profile 16

4. Representative Training Sortie Profile 17

5. Commercial Engine Removal Experience 20

6. Airline Engine Removal Experience vs. OPSEV Model Predictions 23

7. Monthly C-5 Flying Hours and Landings 27

8. Monthly Average Time Between Landings 27

9. Monthly Average Sortie Lengths 28

10. Flying Hours vs. Sorties to Removal 28

11. "Lifetime" Average Sortie Length 29

12. "Lifetime" Average Time Between Landings 30

13. Percentage of Engines Removed by Age Interval 37

14. Cumulative Percentage of Engines Removed 38

15. Conditional Probability of Removal by Age Interval 39

16. Survival Curves for Different Cycle Lengths 40

17. Mean Time Between Removal for Different Cycle Lengths 41

18. Weibull Probability Plot 44

19. Predicted and Observed Survival Curves 45

20. Exponential and Weibull Survival Curves 57

A.1. Stockage Requirements with Varying Demand Rates 72

A.2. Stockage Requirements with Varying Demand Rates and Reduced Pipeline Time 72

TABLES

1. Sample Size Requirements for the Logrank Test 51

A.1. Pipeline Times Used for TF39 Engine Requirements Computations 70

A.2. Effects of Changing Parameter Values 71

B.1. OPSEV Model Inputs 77

B.2. OPSEV Predictions for Other Applications 77

I. INTRODUCTION

The Air Force's fiscal year 1983 budget for procurement of exchangeable aircraft components--assemblies and subassemblies that can be repaired and restored to service after they fail--was more than $3 billion. This figure includes funding for both new items entering the inventory and replenishment of inventory levels for existing items. It is projected to increase greatly through the decade.

Over half of this funding is required to maintain inventory levels used to support peacetime operations, and the balance will be used to procure stocks needed for support of an increased wartime operating tempo. Requirements for "war reserve" material, as well as those for the facilities and equipment needed to accomplish maintenance in wartime, are based on extrapolations of peacetime experience to wartime activity levels.

Wartime requirements forecasts reflect several assumptions concerning the underlying logistics parameters. One particularly critical assumption concerns the relationship between the demand for support resources and the aircraft flying program.

Usually it is assumed that the component demand rate--expressed in demands per flying hour--is a constant that can be:

- Derived from peacetime experience;
- Multiplied times wartime flying hours to yield expected wartime demands.

This Note will demonstrate that the "flying hour" model is too simplistic; the demands per flying hour ratio can be expected to change if aircraft utilization in wartime will differ significantly from peacetime experience. The Note will also show that peacetime operating experience currently does not provide an adequate basis for developing an estimate of the wartime demand rate that could be used to improve war reserve requirements estimates.[1]

[1]A recent Air Force study found that considerable improvement is needed in the methodology used to forecast peacetime requirements (Department of the Air Force, 1983). Forecasting methodologies for wartime deserve even more attention.

STATISTICAL AND POLICY ISSUES THAT AFFECT REQUIREMENTS FORECASTS

From an engineering standpoint, the assumption that demand is strictly a function of aircraft flying hours is very tenuous. Operating time can result in aging that may increase the probability of failure for some components. However, most failure mechanisms are considerably more complex, and the factors that affect component reliability undoubtedly vary across component types. For example, the life of electronic components is probably affected more by the number of power surges to which they are subjected than operating hours; landing gear life should be influenced primarily by takeoffs and landings; and engine life should be affected by rotational speed changes, which are accompanied by heat transients and pressure changes.

Operating time provides an adequate proxy for these factors for peacetime requirements computations because of a peculiar institutional feature of peacetime operations. Because operating patterns do not vary much from one year to the next in peacetime, the ratio of sorties, landings, etc. to flying hours is quite stable over time. Flying hours can therefore serve as a proxy for these other kinds of stresses. In wartime, however, there will be radical changes in the way aircraft are used. Therefore, even if the operating time proxy is adequate for projecting peacetime requirements, it probably is not sufficient for projecting wartime demand. The stockage requirements methodology for wartime must take account of the underlying failure mechanism if it is to generate reasonable estimates of wartime needs.

Aircraft are used in various peacetime missions, but they do not generate the data needed to forecast wartime requirements because key utilization measures, such as sortie length and time between landings, do not vary much across aircraft during a year. The data needed to generate improved wartime forecasts could be generated from peacetime operations but are not because the need to collect useful logistics data is not considered in aircraft scheduling policy.

Currently aircraft are assigned to missions based on training demands and operational commitments, and no effort is made to commit specific aircraft to particular types of missions for extended periods of time. The effect of this scheduling practice is to reduce variability in utilization statistics over time. In fact, in some cases reducing variation in utilization is a scheduling objective.

Attempts to assess the validity of, or improve, wartime support resource requirements estimates must thus confront several interrelated issues:

- The statistical assumptions embedded in current requirements methods have large resource implications;
- The validity of these assumptions has not been established and cannot be tested because varability in peacetime aircraft utilization is lacking;
- Current scheduling policies are unlikely to generate the data needed in the future;
- Some changes are needed in aircraft scheduling policies to provide an empirical basis for requirements forecasts.[2]

This Note uses the current problem of estimating wartime requirements for spare engines for the C-5 as a means to elucidate these issues.

THE C-5 ENGINE REQUIREMENTS PROBLEM

Aircraft engines are the most expensive exchangeable items used by the Air Force. Estimates of the parameters used to compute engine requirements have particularly serious resource implications. If the wartime and peacetime demand rates differ, requirements estimates that reflect a constant demand rate are erroneous.

This issue is both important and timely. The Air Force is in the process of buying 50 C-5B aircraft to augment the existing fleet of 77. This aircraft procurement will generate an additional requirement for spare TF39 engines, as well as the modules and spare parts needed to repair these engines.

[2] Engineering estimates may provide a means to limit the impact of these changes.

Air Force records indicate that there are 150 spare TF39 engines, or nearly one spare per two installed engines, currently in the inventory. Some of these "engines" have been cannibalized[3] extensively and now consist of little more than a nameplate. Nonetheless, if spare engines were procured for the new aircraft at this same ratio, the cost for spare engines alone would approach $300 million. The costs of additional spare engine modules and components would further increase this investment.

Current plans do not include maintaining this ratio of spares to installed engines, but the Air Force does plan to buy another 25 engines. If the assumptions and methodology used to size wartime inventories result in an overstatement of spares requirements, a considerable portion of this more than $100 million cost could be avoided without undermining established wartime capability goals.

The original objective of this research was to develop an improved methodology for estimating wartime logistics parameters and to use this methodology to estimate wartime engine support requirements for the C-5. The hypothesis was that cumulative engine stress could be measured by a combination of engine operating hours and engine "cycles," or large rotational speed changes.

Unfortunately, the lack of variability in peacetime aircraft utilization, coupled with a change in the configuration of the TF39 engine, made it impossible to complete this task. This Note presents an approach to the problem of estimating wartime support resource requirements that would relax the constraints created by limitations of the available data. This approach, which uses a combination of engineering and peacetime maintenance data, is relevant to the specific problem of computing requirements for spare turbine engines used on the C-5 and other aircraft, as well as the more general problem of estimating wartime support resource requirements.

---

[3]"Cannibalization" is the use of parts from one engine to repair another. This maintenance practice consolidates spare parts shortages to minimize the number of engines out of service. The missing parts should eventually be replaced, but some TF39 engines have been cannibalized so extensively that it might be cheaper to purchase new ones than to replace all of the missing parts.

OUTLINE OF THE REPORT

Section II describes the parameters used in spares requirements computations, some of the weaknesses of the assumptions on which parameter estimates are based, and the computational methodology that uses these estimates. Section III contains a discussion of the engineering reasons for engine performance deterioration or failure and summarizes some of the lessons learned from commercial experience with aircraft turbine engines. This experience is reflected in an engineering model that has been used to predict engine removal rates in a variety of applications. Both airline experience and the model's predictions are consistent with engineering explanations of the reasons for engine removal.

Engineering theory and commercial experience, however, do not provide an adequate basis for projecting the removal rate for C-5 engines in wartime. Differences between the engines used in commercial and military applications, compounded by differences in both utilization and maintenance practices, preclude direct application of commercial data to the military requirements problem. Section IV therefore describes the Air Force data that were collected to estimate the wartime C-5 engine demand rate as well as an exploratory data analysis that established the inadequacy of available peacetime data for this purpose.

Section V describes an alternative approach to the data analysis problem--application of actuarial, or "life" data, analytic techniques. Although the lack of variability in peacetime utilization data also constrains this approach, the results of the analysis further reinforce the hypothesis that aircraft utilization affects the removal rate. This analysis also provides an introduction to some of the issues addressed in the discussion of a proposal for generating more useful data presented in Sec. VI.

Section VII addresses the implications of the proposal for peacetime C-5 operating policies. It also describes some potential applications of the analytic approach to other aircraft systems and its implications for a wide range of support resource requirements.

II. ESTIMATING STOCKAGE REQUIREMENTS

Most stockage requirements methodologies used within the Department of Defense assume that demand is generated by a simple Poisson process.[1] This assumption implies that the probability of having a needed part on hand when it is required, or the probability of no "stockout" [P(NO)], is given by the expression:

$$P(NO) = \sum_{k=0}^{s-1} \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

where

λ = The daily demand rate for the item

t = The time needed to repair and return an item to the operating forces

k = The number of components tied up in repair or transportation "pipelines"

s = The stock level maintained for the part.

This formula is used to compute both peacetime and wartime requirements. P(NO) for engines is set at .8 as a matter of policy--engine "requirements" provide 80 percent confidence that an engine will be available when needed--if the demand process is simple Poisson with parameter λ.[2]

---

[1]The *events* in a simple Poisson process occur singly and are statistically independent; the occurrence of one event does not affect the probability of another event. This is a mathematical description of a process in which the events are totally random, hence completely unpredictable. The interarrival times for a Poisson process are distributed exponentially, and the only parameter that must be specified to describe it is the average number of events per period (e.g., flying hour, day) (Walker, Chaiken, and Ignall, 1979). It is reasonable to assume that the removal process is simple Poisson unless there is compelling evidence to contradict this assumption (conversation with William Rogers, a statistician at The Rand Corporation, 1980).

[2]If this formula is applied to each item individually, and the probability of stockout for each item is set at the same level, the resulting stock level will provide approximately even protection against

The assumption that underlies both peacetime and wartime aircraft component stockage requirements estimates is that part removals, hence demands for stock, are driven by the aircraft flying hour program.[3] Peacetime demand experience is converted to a demand rate by dividing the number of requests for an item (demand) by the number of hours flown during the period. This rate of demands per flying-hour is then converted to peacetime and wartime Daily Demand Rates (DDRs) by multiplying it by the projected peacetime and wartime daily flying programs.

The DDR is used in conjunction with estimates of the length of resupply pipelines--the amount of time needed to move, accumulate, repair, and return failed components to the aircraft--to determine stockage requirements. The pipeline time used for computing requirements is a weighted average of the time spent at different maintenance echelons and in the transportation segments connecting the other elements of the support system to the flight line.

The logistics system described above is shown graphically in Fig. 1. Demands for parts to replace items removed at the flight line "pull" stock from the supply system, and failed components are "evacuated" for maintenance. Components that can be repaired locally are returned to local supply to satisfy future demands. Those beyond the capability of

---

stockouts across the inventory. Requirements methodologies for some resources use marginal analysis techniques to provide a specified level of protection across the inventory rather than a fixed level of protection for each item.

[3]As was suggested above, one of the key parameters used in requirements computations is the demand rate, usually expressed in demands/flying hour. It is useful to distinguish among (1) failures, (2) removals, and (3) demands. A *failure* is defined as an unsatisfactory condition. It may be either a *functional failure*, which is an inability to meet a specified performance standard, or a *potential failure*, which is an identifiable physical condition indicating that functional failure is imminent (Nowlan and Heap, 1974). Failures lead to *removals*, and usually also generate *demands*. However, in peacetime not all removals reflect failure, nor do they necessarily generate demand. *Demand* is the event of interest for requirements computation because it levies a requirement on the logistics system. The demand rate will exceed the functional failure rate because potential failures (and even some functional failures) are defined by policy. The peacetime component demand rate is less than or equal to the removal rate; the demand rate should equal the removal rate in wartime.

local maintenance, or "Not Repairable this Station" (NRTS), are shipped to the depot to be repaired and returned to central system stocks. They are transferred back to a base when local inventories are depleted, or when demands cannot be satisfied from base supply.

Pipeline stock, or the expected value of pipeline contents, is equal to the product of DDR and pipeline time. If spares are not provided to cover these pipeline requirements, the necessary components will be cannibalized from other aircraft, and some aircraft will be unable to perform their missions because parts are lacking.

Spares computations also provide protection against stochastic variability in demand and resupply times with "safety levels." The total stock level is set to provide a specified level of confidence that a component will be available when needed, or to minimize backorders subject to a cost constraint, on the assumption that the demand process is Poisson.

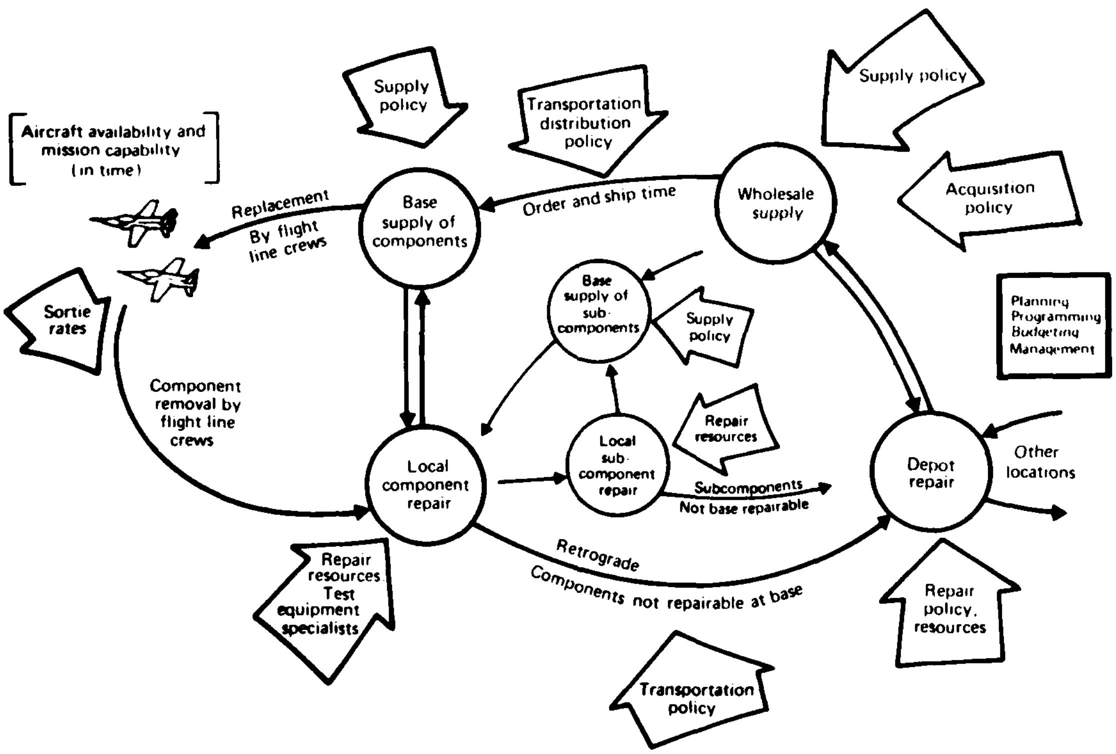


Fig. 1 — The aviation logistics support system

In summary, requirements computations produce stock levels that are intended both to cover pipelines and provide protection against stockouts due to variation in pipeline contents. Peacetime experience is normally the basis for estimating both the DDR and pipeline times.

Spares requirements are particularly sensitive to the demand rate used in computing requirements. The relative influence of the key logistics parameters on wartime TF-39 engine requirements can be seen in Fig. 2,[4] which shows the percentage change in spares requirements as each of the parameters is varied by a specified percentage when:

- The flying program is fixed, and
- All other parameters are held constant.

The influence of the demand rate is to be expected, because the daily demand rate is multiplied by each of the pipeline segments to develop an estimate of the total pipeline requirement. The relative effects of the parameters that determine pipeline length would differ for other components, but the demand rate would remain the most influential factor.

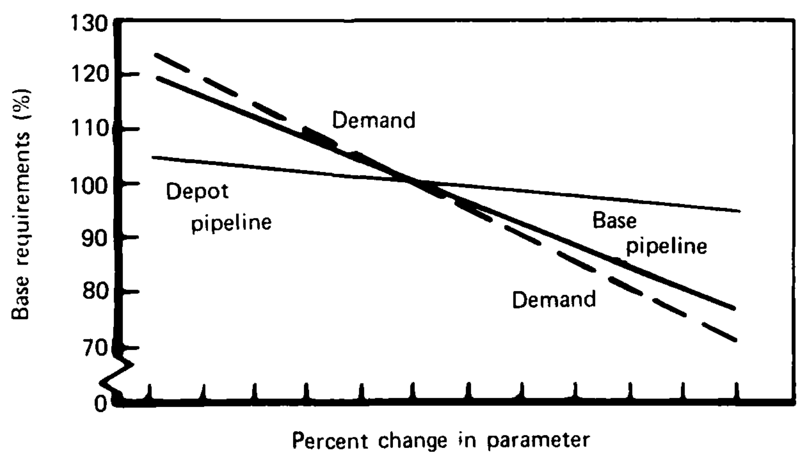


Fig. 2 — Requirements sensitivity to parameter estimates

[4]The lines plotted are continuous to simplify comparison. In reality, stock can be procured only in integer quantities, so the change in stockage requirements is a step function. Further examples of the effects of changing parameter values are provided in Appendix A.

WEAKNESSES IN THE ASSUMPTIONS USED TO DEVELOP STOCKAGE REQUIREMENTS

Several past studies have cast doubt on the assumption that component demand is (or should be) proportional to the flying program. This assumption is often incorrectly labeled the "linearity" assumption, even though proportionality is only one of an infinite number of possible linear relationships between flying activity and demand.

Most of these studies have dealt with the effects of changing sortie lengths (e.g., Casey, 1977; Donaldson and Sweetland, 1968; Hunsaker, 1977; Kern and Darns, 1966; Sallee, 1974; Shaw, 1982;, Shurman, 1970), although some have addressed the effects of increasing aircraft utilization rates with constant sortie lengths (e.g., Embry and Crawford, 1983). The majority of this work has been based on data analysis rather than tests of engineering hypotheses, but the data are consistent with engineering explanations of the variable failure rate phenomenon.

These studies suggest that the linearity assumption may lead to an overstatement of spare TF39 engine requirements. The wartime employment of strategic airlift aircraft will differ considerably from peacetime experience. Utilization rates (in flying hours per day) will increase by a factor of five to six, the sortie mix will shift to favor heavily loaded cargo flights rather than lightly loaded training missions, average sortie lengths will increase, and the average number of landings per sortie will decrease greatly. All of these factors can be expected to change the demand for engines from that experienced in peacetime.

Although this study emphasizes the problem of estimating the demand rate, the assumption that wartime pipeline times will approximate those experienced in peacetime is also suspect.[5] The time needed to repair components in wartime could decrease because personnel and facilities are utilized more intensively, or might increase if capacity constraints cause queuing in the maintenance process. Transportation times can also be expected to change under wartime conditions. Furthermore, maintenance policy would probably change in response to wartime demands.

Operational realities would change perceptions concerning the need for maintenance, as well as the availability of resources to accomplish it.

Finally, the assumption that demand is generated by a simple Poisson process is questionable. Various data analyses conducted over the years have found that the variance-to-mean ratio of demand often exceeds that implied by the Poisson assumption and that removals of some components are not independent events.[6]

The exponential distribution of service life is characteristic of Poisson models (Tribus, 1967), and implies that the probability of failure is independent of age. In contrast, the Weibull distribution allows recognition of a time-dependent failure rate--an increasing (or decreasing) probability of failure over time. Several analysts have concluded that distributions of "lifetime" (time to failure) data for complex, multi-component items, including aircraft turbine engines, fit a Weibull distribution (e.g., Barlow and Proschan, 1975; Chorafas, 1960; Walker, 1980).[7]

The level of wartime capability attainable for a fixed defense budget is constrained unnecessarily to the extent that any of these assumptions result in an overstatement of wartime resource requirements. Overinvestment in any one resource category consumes resources that

---

[5]In some instances, wartime pipeline times are established by policy. For example, spares kits for tactical aircraft have been computed assuming that local repairs can be accomplished in two days. This would be a difficult goal to achieve even if the necessary maintenance resources were available, but logistics plans do not provide for delivery of these resources until later in the scenario.

[6]The distribution of a simple Poisson process has a variance that is equal to its mean. Variance-to-mean ratios greater than one have been observed in logistics data for several years (e.g., see Feeney, Peterson and Sherbrooke, 1963). A ratio greater than unity would result if removals were not truly independent events. Because troubleshooting procedures frequently provide ambiguous results, particularly for electronic components, removals may not be independent. An experienced technician may decide to replace a component whether or not the component has failed because replacing it corrected similar problems in the past. Thus a removal may reflect the technician's experience, and the probability of removal may be affected by this experience.

[7]The exponential is a special case of the Weibull distribution. The more general distribution has also been used to describe other phenomena (Berretoni, 1962; Weibull, 1951). It is discussed further in Sec. V.

could be used more efficiently in other areas (Hitch and McKean, 1960). These assumptions may lead to an understatement of requirements for some components, particularly those used on tactical aircraft that will be subject to greater stress in wartime than they are in peacetime. All of these assumptions, particularly those affecting wartime demand rate forecasts, deserve careful attention.

WARTIME FAILURE RATES OF AIRCRAFT ENGINES

Several approaches could be taken to the problem of estimating wartime turbine engine removal rates. These alternatives include:

- Utilizing data obtained by operators of similar engines (e.g., the airlines) to account for the effects of changing utilization patterns;
- Applying engineering models of the failure process to project removal rates as a function of utilization;
- Employing peacetime data to statistically estimate wartime parameters;
- A combination of the above.

Each of these estimating methodologies is open to criticism. For example, military operators are quick to point out the differences between civilian and military operations; engineering models have never been considered to be a reliable source of removal rate estimates; peacetime data obviously do not reflect wartime conditions; and any combination of the alternatives will inherit the deficiencies, as well as the strengths, of the individual approaches considered.

Nonetheless, a removal rate estimate is needed to develop war reserve stockage requirements. Section III presents an engineering rationale for expecting a change in the engine demand rate in wartime, summarizes airline experience that supports the engineering hypotheses, and describes an engineering model that captures the effects of this commercial experience. Analyses of available military data are provided in Secs. IV and V.

III. AIRCRAFT TURBINE ENGINE RELIABILITY

Turbine engines have been installed on most commercial and military aircraft manufactured in the last two decades. Considerable effort has been dedicated to understanding and improving the reliability characteristics of these engines for both safety and economic reasons. The subjects of this research have ranged from metallurgy and fluid mechanics to development of monitoring and diagnostic systems to track engine operating parameters through the various stages of flight. Analysis of the data generated across the research spectrum has provided engineers with a reasonably good understanding of what causes engine failures and an ability to estimate engine removals as a function of utilization.

The research has not progressed to the point that engineers can predict with any certainty when a particular engine will be removed, partly because there are many reasons for engine removal, and engine failure is a random process. Dominant failure modes have been identified, however, and these modes have been associated with parameters that vary with engine utilization. This section discusses the effects of utilization on observed engine reliability and its implications for development of estimates of the wartime engine demand rate for strategic airlift aircraft.

REASONS FOR ENGINE REMOVAL

Engine removal reasons can be divided into two broad classes:

- "Engine-caused" attributable to
  - Failure of an engine component or
  - Deterioration in engine performance
- Attributable to other causes, such as
  - Foreign object damage (FOD)
  - Other environmental factors
  - Scheduled (preventive) engine maintenance
  - To facilitate other aircraft maintenance
  - Cannibalization.

Some FOD results from carelessness, such as leaving a wrench in the engine, but a much more common cause is ingestion of material from the runway or a bird while in flight. Thus the FOD rate should be influenced by the frequency with which the aircraft enters an environment in which it is at risk to foreign object damage. Other environmental reasons for engine removal are associated with the operational environment.[1] Presumably most removals that are not engine-caused could be postponed during wartime.

Many engine caused removals occur when operating parameters exceed acceptable limits. For example, an engine may be removed because it has insufficient turbine inlet temperature (TIT) margin.[2]

Exhaust gas temperature (EGT) is also watched closely because it is an indicator of internal wear that reduces engine efficiency. The TIT margin will eventually deteriorate, or the EGT rise, to the point that the engine must be removed to restore operating tolerances even if no component failure has occurred.[3]

Many engines do not "live" to the point that removal will be required simply to restore internal tolerances. A failure of one or more of the engine's internal components is likely to cause a removal for corrective maintenance before the engine has an opportunity to wear

---

[1]For example, sand may be drawn through the engine during takeoffs and landings in a desert environment, which would increase component wear. The effects of such adverse environmental conditions can be estimated for different operational scenarios.

[2]"TIT margin" is the difference between observed and allowed turbine inlet temperature. Engine performance (thrust) can be improved by increasing turbine inlet temperature, but only at the expense of the engine's useful life. "Trimming" to increase thrust raises temperature in the engine and may lead to premature failure.

[3]The number of removals for reasons other than component failure may also be affected by the frequency of exposure to maintenance, the level of workload in maintenance facilities, etc. The maintenance activity can exercise some latitude in scheduling a removal when performance parameters are beyond or near established limits. Hence it has been suggested that the demand rate may be in part a reflection of the workload in maintenance facilities rather than simply a measure of the amount of workload than can be expected to be generated.

out. Failure of one component generally leads to collateral damage in other parts of the engine, so determining the reason for failure does not itself identify which components will have to be replaced. An understanding of primary failure modes, however, can be used to predict the number of failures likely to be experienced over some future time. Further analysis of primary and secondary failure effects can then be used to estimate parts requirements.

A primary cause of component failure is a phenomenon known as "low cycle fatigue," which is usually associated with turbine rotor speed changes. Additional stress is created by the large heat transients and pressure changes that occur during takeoff and landing. The number and duration of such stressful events will influence component, hence engine, life. This phenomenon is not unique to aircraft turbine engines; low cycle fatigue is a major concern in other applications where metal expands and contracts repeatedly. For example, changes in temperature and pressure in steam generators and piping systems used in electrical generating plants can result in low cycle fatigue failures.[4]

Operating hours are not a good proxy for the stress due to low cycle fatigue, although they may be important to the life of mechanical components (e.g., bearings). Operating time may also affect the life of other components. Since failures may be related to hours, cycles, or both, the problem of estimating "engine caused" failures is one of:

- Identifying engine failure modes;
- Associating appropriate measures of engine utilization with these modes.

MEASURING ENGINE STRESS

The operating profile for an aircraft turbine engine employed in a typical (commercial or military) point-to-point sortie is displayed in Fig. 3. Starting the engine causes a temperature transient that continues through the takeoff and climb-out phases of flight; its intensity and duration is a function of the load the aircraft is

[4]The definition of low cycle fatigue in these other applications is obviously broader than that used for describing turbine engine stress. An engineer from San Diego Electric pointed out the importance of fatigue effects during a conversation with the author during 1983.

carrying, the length of the runway from which the sortie originates, and the atmospheric conditions at the airport or base. Temperature and speed then fall for the cruise portion of the flight, after which they rise again while the pilot makes preparations for landing and then reverses thrust to brake the aircraft on the ground. This ambient-to-ambient temperature profile is commonly counted as an engine "cycle."

The speed and temperature profiles for commercial airliners and military cargo flights, are considerably different from those of tactical aircraft. A fighter sortie would include numerous throttle changes to adjust for maneuvers and changes in the tactical environment. These throttle changes also create cyclic stress, and a fighter sortie would produce more of this type of stress than a cargo-hauling sortie.

Many peacetime strategic airlift missions are also considerably more stressful than Fig. 3 would suggest. The peacetime mission of the strategic airlift force is training (to prepare pilots for their wartime missions), and many sorties include multiple landings. A representative profile for a training sortie is shown in Fig. 4.

Although the stress on the engine resulting from "excess" landings is not as great as that created by the increase in speed and temperature that occur when the engine is started, a sortie with more than one

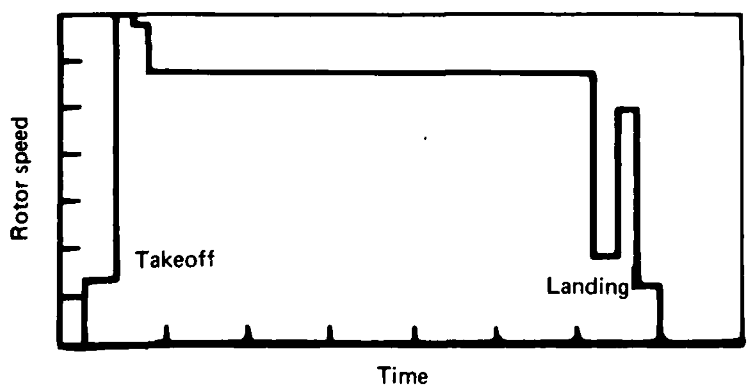


Fig. 3 — Typical mission sortie profile

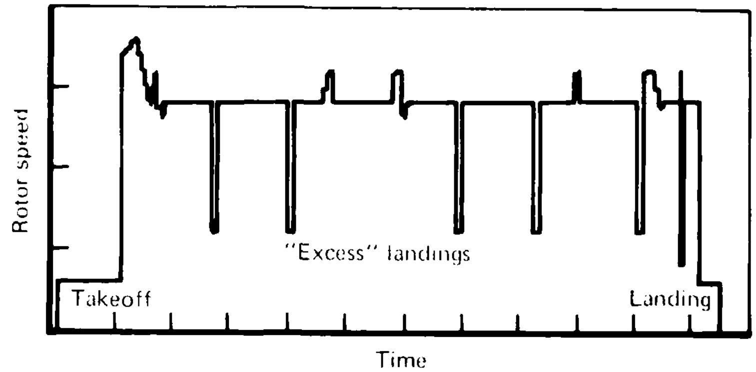


Fig. 4 — Representative training sortie profile

landing is harder on an engine than a point-to-point sortie.[5] In fact, differences between the planned and actual engine "duty cycle" have been shown to greatly affect engine deterioration and removal experience (Birkler and Nelson, 1979; Cote and Birkler, 1979).

Cycle accrual is one of the measures recorded in the C-5 Ground Processing System (GPS), a data system that records TF39 utilization statistics (Department of the Air Force, 1981). Training sorties are counted as two "cycles" in this system. Past studies of airline data have treated within-mission throttle changes as half-cycles (Byrd, Barrial, and Kostomay, 1978).

IMPLICATIONS OF THE MISSION PROFILE FOR THE ENGINE FAILURE RATE

Engine life is affected by both environmental factors and utilization. Based on their understanding of the materials used to build engines, their analysis of the data accumulated using engine monitoring and diagnostic systems, and detailed analysis of both failed engines and statistical failure data, engine manufacturers have identified three primary utilization measures that influence engine life:

[5]Training sorties may also create the conditions for a phenomenon known as "hot rotor reburst." The static elements of an engine cool faster than rotating elements. Differences in the rate of expansion and contraction of different engine parts that result from cutting power to idle and then reapplying it can cause rotor blades to "chew out" the turbine case (Wulf, 1980).

- Cycle accrual;
- Operating time;
- Time at maximum operating temperature.

Most of the engine monitoring systems that have been designed to track engine performance deterioration record measures of these parameters (see National Aeronautics and Space Administration, 1981).

As is suggested by Figs. 3 and 4, strategic airlift engines will be subjected to fewer cycles per sortie in wartime than they are in peacetime. Engine "cycle density," or cycles per flying hour, will decrease from a peacetime value of about .38 to about .2 in wartime. Engines may experience longer operating times at maximum temperature in wartime because the aircraft will be carrying heavier loads and may have to operate from shorter runways. The extent to which that effect will offset one attributable to a lower cycle per flying hour ratio has not been determined.[6] However, even in wartime most sorties will not be flown fully loaded. On balance, the changes in wartime flying should result in a reduction in the engine-caused removal rate.

Even though the probability of FOD might increase in wartime, the frequency of exposure to FOD should decrease with decreasing landings per sortie, so FOD removals per flying hour should not change very much. Finally, most convenience and scheduled maintenance removals can and should be postponed during the surge period. The net result is likely to be a reduction in the engine demand rate observed during wartime.

---

[6]The engine manufacturer computes the percentage of total operating time spent at full throttle for a reference and other missions and uses the ratio of this percentage to estimate the difference between expected removal rates for different mission profiles. Although the time spent at maximum throttle on any given wartime sortie may exceed that for peacetime cargo missions, the percentage of total operating time spent at full throttle should not increase much because average sortie lengths will increase in wartime.

AIRLINE ENGINE REMOVAL EXPERIENCE

Turbine engines power the modern jet aircraft used by all of the major airlines. Because propulsion costs represent a considerable fraction of their total operating costs, airline operators are very interested in understanding the reasons for engine failure, as well as the relationship between failure rate and operating parameters. The diversity in route structures within and across airlines also ensures considerable variability in utilization data. Hence airline experience is relevant to the military demand rate estimation problem, even if the results of this experience cannot be used directly to compute military resource requirements.

Engine and aircraft manufacturers have an even stronger motivation to understand the factors that drive engine requirements. Their customers consider operational reliability an important performance parameter and may require warranties as a condition of purchase.[7] The major engine manufacturers spend a great deal of time and money collecting engine utilization and failure data from all of their customers, including the military. These data are used for improvement of existing products and as an input to new designs.

Airline maintenance executives and analysts have long been aware that maintenance requirements vary with aircraft utilization. Commercial operators are driven to achieve utilization rates of 12 or more hours per day, not only because economics dictate that expensive capital equipment should be fully utilized, but because maintenance demands per flying hour have been observed to decline with increasing utilization rates.[8]

---

[7]The military services have also obtained warranties on some of their equipment purchases, but experience with such warranties has been mixed (Gandara and Rich, 1977). Nevertheless, the Congress has legislated a requirement that most future major procurements include warranty provisions.

[8]Conversation with the former Vice President for Maintenance, United Airlines, May 1983. The specific example cited was the Caravelle jet. Maintenance costs became intolerable when Caravelle utilization rates were descreased to less than eight hours per day.

Figure 5 shows the removal rates for CF6 engines used by three groups of airlines. Operators with the longest average sortie length experience the lowest removal rate. Conversely, airlines with shorter average sortie lengths have higher removal rates (Byrd, Barrial, and Kostomay, 1978). Aircraft engine removal data are difficult to compare across airlines because of differences in maintenance policies and practices, the data presented in the figure reinforce the hypothesis that removal rates decrease with increasing sortie lengths. (Several related commercial studies are American Airlines, 1967; Carlyle, 1976; Davis, 1967; Sallee, 1974; Wulf, 1980).

The CF6 engine used in commercial service is a derivative of the TF39 that is installed on C-5 aircraft. There are some differences in both the manufacturing and operating characteristics of these two engines. Nevertheless, the tendency for engine removal rates to decline with increasing average sortie lengths noted by commercial operators can be expected to hold for the C-5.

Airline sortie length data are equivalent to cycle data because the airlines seldom perform multiple landings during a single sortie. Average C-5 sortie lengths are expected to increase in wartime, and average cycle lengths would increase even if sortie lengths were held constant because only one landing will be made on most sorties (vs. a

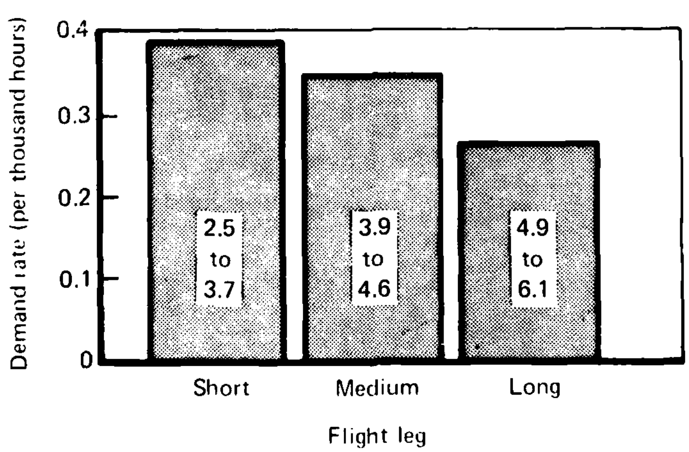


Fig. 5 — Commercial engine removal experience

peacetime average of 3.07). This implies that the wartime TF39 removal rate should be lower than that experienced in peacetime.

Unfortunately, the airline data cannot be used to infer how much of a reduction in TF39 removal rates should be observed in wartime. Thus these data merely reinforce the hypotheses that could not be tested using peacetime C-5 flying data. However, they also encourage further exploration of alternatives for developing removal rate estimates.

AN ENGINEERING MODEL OF THE ENGINE FAILURE PROCESS

General Electric has developed an engineering model of the engine failure process that reflects the effects of the three types of stress identified above (Byrd, Barrial, and Kostomay, 1973). It has been used to predict removal rates under various operating conditions, including use of the core engine in ship propulsion and pump operating applications.

The model estimates the failure rate for a particular mission by comparing the "severity" of the mission to a "reference" mission. "Severity" is defined as the ratio of failure rates; the objective of the model is to estimate this ratio based on detailed analysis of component failure modes and the mission profiles of the "reference" and "comparison" missions.

The severity ratio is developed as follows:

- Subdivide the engine into major assemblies;
- Establish reference failure rates for each major assembly based on:
  - Comparison with similar systems
  - Failure modes and effects analysis (FMEA)
- Define a reference mission
- Specify comparative missions
- Develop utilization measures (e.g., cycle accrual) for these missions and estimate their effects on component removals

- Compute severity ratios by dividing the estimated failure rates for the comparative missions by that for the reference mission.

The key to the analysis is estimation of the relative contributions of cyclic and hour effects to the failure rates of the individual components. These estimates depend on engineering judgment, as it is difficult to ascribe failures to a particular type of stress through either examination of failed engines or data analysis. Hence the model inputs cannot be verified analytically; the only measure of their validity is the accuracy of the model's predictions.

The model's predictions, when compared with actual airline operating experience, have proven to be remarkably accurate. Figure 6 the reciprocal of the observed demand rates shown in Fig. 5, or the Mean Time Between Removal (MTBR), in contrast to the predicted values for the same parameter. The predictions appear to be conservative, and this tendency increases with increasing flight leg length. The fact that the predicted removal rates are consistently higher than those observed by the airlines should allay some Air Force concerns about using this model. Appendix B provides an additional description of the model and representative inputs.

The inputs to the model suggest that over half of expected reference mission failures are attributable to cycle-related stress if time at maximum operating temperature is held constant. Increasing cycle density would further increase the relative contributions of cycles to the total failure rate. This is consistent with the results of Wulf (1980), which found that both engine cycles and operating time contribute to engine deterioration. Increasing the time at maximum operating temperature is predicted to increase the failure rate due to both cycle- and operating time-induced stress.

The parameter values that are needed to drive the engineering model for the TF39 would differ from those used for the CF6. To apply the model to the problem of estimating C-5 wartime engine removal rates would require an "up front" investment, but it would be small relative to the investment in spare engines the Air Force now contemplates. The potentially high payoff of applying this model suggests that it would be extremely worthwhile.

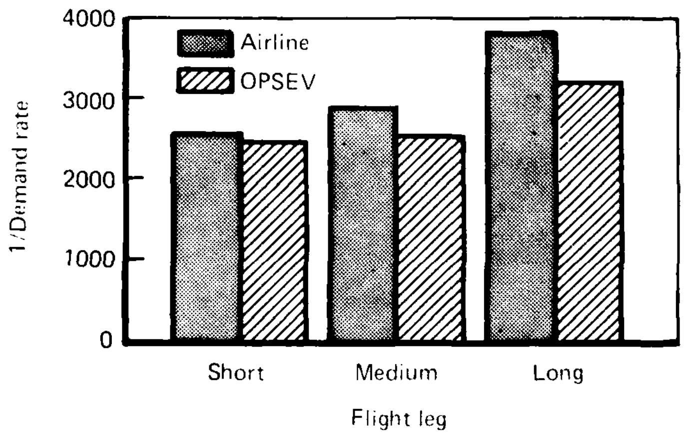


Fig. 6 — Airline engine removal experience vs. OPSEV model predictions

IV. C-5 ENGINE UTILIZATION DATA AND EXPLORATORY DATA ANALYSIS

The engineering rationale and airline experience presented in the preceding sections reinforce the hypothesis that MTBR increases with decreasing cycle density. Models for projecting the wartime C-5 engine removal rate, however, should be based on military data.

This section describes the TF39 utilization data that were obtained to support statistical modeling of the C-5 engine removal process and summarizes the data processing procedures.[1] An exploratory data analysis exposed an important limitation of the data: a lack of variability in peacetime aircraft utilization, which is attributable to aircraft scheduling policies designed to equalize utilization across aircraft.

DATA DESCRIPTION

The primary source of the data extracted for the analysis was the Ground Processing System (GPS), a special data collection system developed for the C-5 that receives input from on-board monitoring instruments and ground personnel (Department of the Air Force, 1981). Compared with the data available from other Air Force data systems, GPS data are fairly clean. Unfortunately, the removal reasons specified in the relevant Air Force Technical Orders (TOs) generally do not provide a reliable indicator of engine failure mode, so the failure mode cannot be determined from the engine histories developed from GPS data files.

Extracts from two files maintained in the GPS system were used:

- Aircraft flying histories, which record daily sorties, landings, and flying hours by aircraft tail number
- Engine installation and removal histories, including removal reason, by engine serial and aircraft tail number.

---

[1] A more detailed description of the data processing is provided in Appendix C.

The engine utilization and engine history records were merged by aircraft tail number to create two engine data files. These files contain:

- Detailed utilization records for each month an engine was installed
- Summary utilization statistics cumulated across engine "lifetimes."

Both files contain the sortie, landing, and flying hour measures described above for each engine installed since 1976.

Engine lifetimes as defined in the second file differ from the normal Air Force definition in one important respect. The Air Force "ages" engines from first delivery or last overhaul. Engine hours are accumulated for the entire period between overhauls, regardless of the number of times the engine is removed for maintenance. The lifetimes defined in the dataset prepared for this analysis date from the engine's last visit to a maintenance shop. The implicit assumption is that all engines having undergone any maintenance--in the depot or a field maintenance activity--have an equal chance of "surviving" future utilization. The alternative assumption, which would have required separating the engines into at least two groups, did not appear to be warranted. Engines fresh from the depot were at least as likely to be removed in the first hundred hours as engines produced by a Jet Engine Intermediate Maintenance (JEIM) shop.

EXPLORATORY DATA ANALYSIS

A primary objective of this research was to explain differences in engine lifetimes with variables that describe utilization between installation and removal for cause. Before any attempt was made to fit a statistical model, an exploratory data analysis identified data errors and generally characterized the distributions of engine life, utilization measures, and repair times.[2] Plots of monthly flying activity show a great deal of variability in aircraft average sortie lengths and time between landings during peacetime. This variability can be seen in Fig. 7, which is a plot of monthly landings vs. flying hours for all of the records contained in the working data file.

As might be expected, the figure indicates that some aircraft are used predominantly for long cargo-hauling missions and others are used primarily for multiple-landing training sorties during any given month. These two extremes define a "wedge" in the plot that contains all of the remaining data points. The left side of this wedge shows the maximum sortie lengths for peacetime cargo missions, and the lower edge represents the minimum time needed to circle the field between training landings. The conclusion that some aircraft are dedicated to particular types of missions during a month is reinforced by the frequency distribution of average time between landings presented in Fig. 8.

Figure 9 shows the distribution of average sortie length for these monthly data. It indicates that most peacetime sorties last from 4 to 5.5 hours. The distribution of time between landings suggests that many of these sorties include multiple landings. These averages, however, obscure the true variability in peacetime sortie lengths. A frequency plot for individual sorties would show greater dispersion than can be seen in the monthly data.

Even this month-to-month variability is smoothed out over an engine lifetime. The lack of variability in the lifetime data can be seen in the plot of sorties vs. flying hours between removals presented as Fig. 10. The line plotted in the upper right quadrant of the figure shows the results of a linear regression of flying hours on sorties; the remainder of the line is covered by the individual data points. The tight clustering of points around the regression line indicates that there is little difference in lifetime average sortie length across engines in peacetime.

The frequency distribution of average sortie lengths shown in Fig. 11 confirms this observation. The distribution is much tighter than that of the monthly data presented in Fig. 8.

Figure 12 shows the distribution of average time between landings over installed engine lifetimes. These data are more variable than the sortie data shown in Figs. 10 and 11 and are distributed differently than similar monthly data.

---

[2]The techniques proposed by Tukey (1977) were used in the initial stages of this analysis.

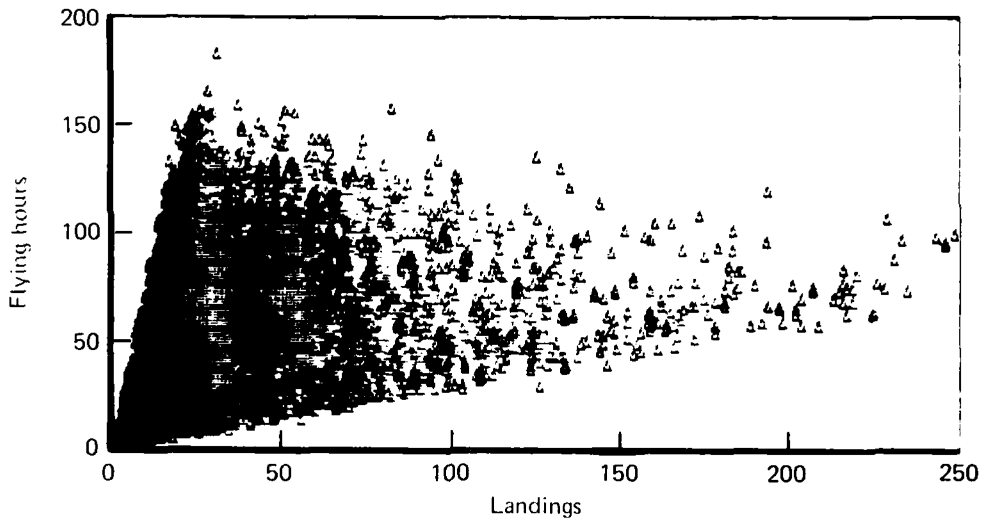


Fig. 7 – Monthly C-5 flying hours and landings

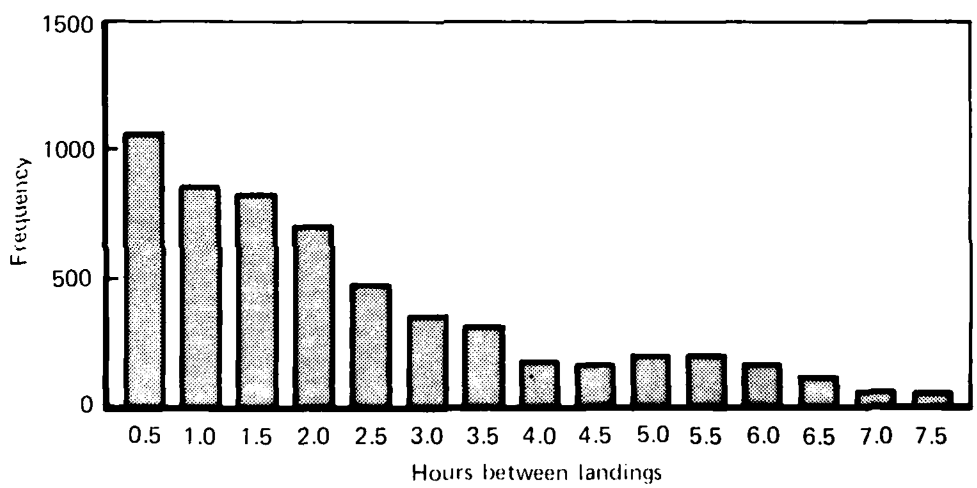


Fig. 8 – Monthly average time between landings

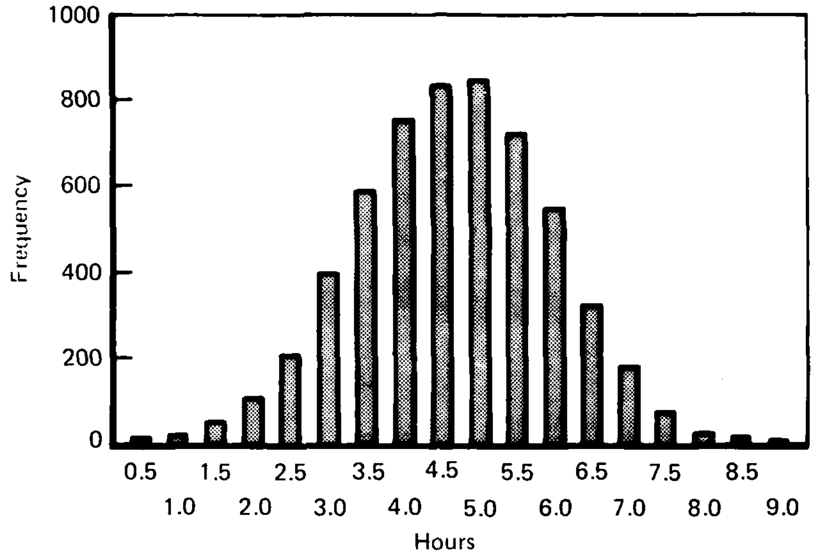


Fig. 9 — Monthly average sortie lengths

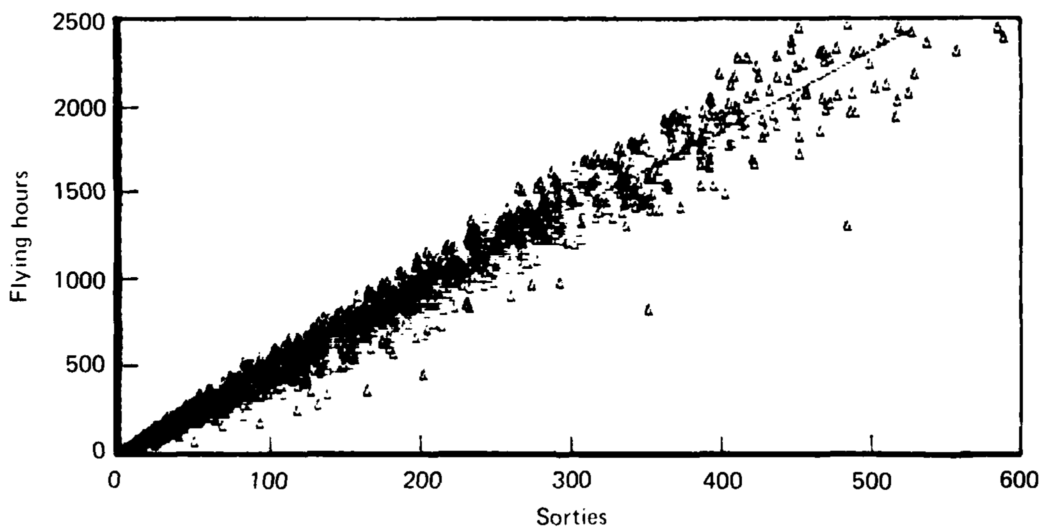


Fig. 10 — Flying hours vs. sorties to removal

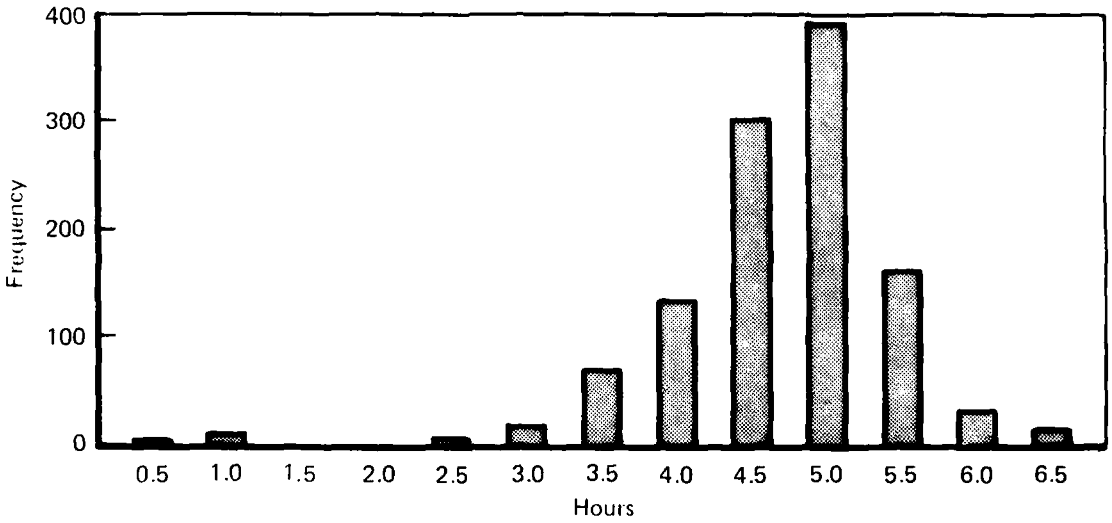


Fig. 11 — "Lifetime" average sortie length

The shift in the modal time between landings and the tightening of the sortie length distribution demonstrate the effect of an aircraft scheduling policy that is intended to equalize the time allocated to different missions across aircraft.[3] The central limit theorem suggests that variability in the sortie length distribution would decrease as the length of time considered increases even if this assignment policy was not being used, but the scheduling policy accentuates the effect.

Air Force success in implementing this scheduling policy is the primary reason that data generated by normal peacetime flying lacks the

---

[3]Engines with short lives account for the few large deviations from the central values in both the sortie and landing data.

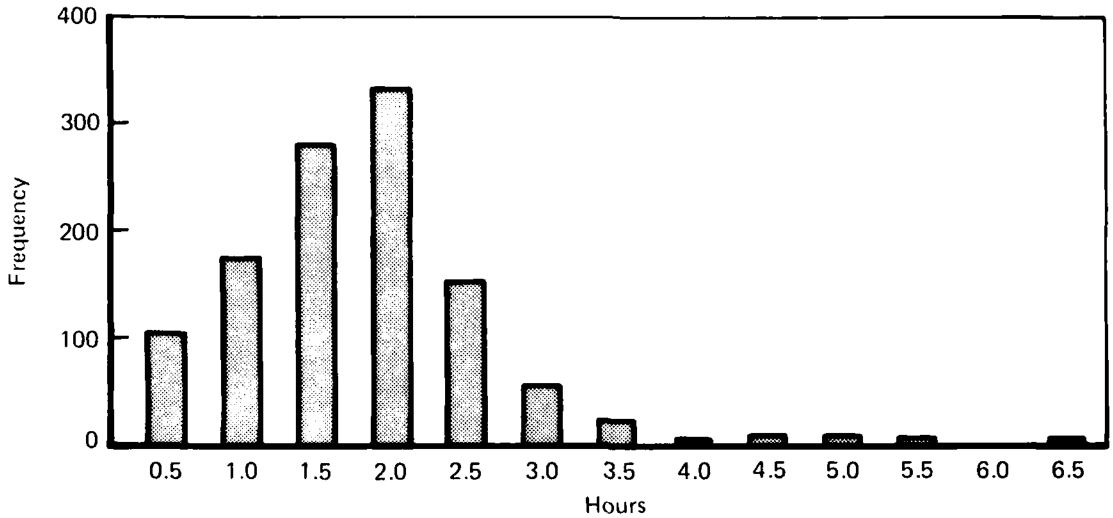


Fig. 12 — "Lifetime" average time between landings

variability needed to develop a model for projecting the wartime removal rate. The natural variation in engine utilization that could be expected if aircraft were dedicated to particular types of missions for extended periods does not exist, and data at the expected wartime average time between landings are not available.

Even though landing data are more variable than sortie data, they cannot be used for testing the linearity assumption because the landings are not necessarily comparable. As noted above, sorties including many touch-and-go landings are more stressful than point-to-point sorties, but the average training landing causes less engine stress than a full-stop landing, particularly of a loaded cargo flight.

The lack of variability in peacetime utilization data constrained statistical analyses intended to provide a basis for projecting the wartime TF39 engine removal rate. The following section provides the results of these analyses. Although the results are not conclusive because of the limitations in the data, they are consistent with those described in Sec. III.

V. ESTIMATING WARTIME REMOVAL RATES

If the engine failure process is more complex than suggested by the simple flying hour model, it should be possible to explain much of the variability in observed engine life with variables that describe engine use during the installed lifetime. The analytic model structured to capture the effects of these other variables could take several forms:

- The failure process is purely random; the mean time to removal for the population is a random variable that can be estimated from the data,
- Failure results from a discrete event immediately preceding or coincident with the failure,
- Failure is the result of cumulative stresses on the engine during its lifetime.

Although each model may apply to some failure modes, none adequately describes the engine failure process. If flying hours between removals are used as the measure of utilization, the first of these alternatives could be no more than a restatement of the linearity assumption. Flying hours do contribute to TIT margin deterioration, which is at least partially an aging phenomenon, as well as bearing wear. The second model describes some catastrophic failure modes, particularly FOD. Finally, past research and engineering theory suggest that cumulative cycles and average cycle length are important measures of stress. They should help to explain unscheduled removals not attributable to FOD. Testing the latter type of model was a primary objective of this research.

Not all of the variability in observed engine life can be explained with variables other than flying hours because some failure modes depend partly on operating age. Identification of cycle-related variables that help to explain variations in engine life would validate the general hypothesis that the flying hour model is too simplistic and would provide a partial basis for predicting wartime removal rates. The data described in the previous section were used in two types of analyses designed to quantify the effects of such explanatory variables.

STATISTICAL ANALYSIS OF ENGINE UTILIZATION DATA

The objectives of the first of the two approaches to analysis of the engine data were to:

- Establish whether engine experience just before the removal explained engine failure,
- Develop a model of the removal process using utilization parameters cumulated over an engine's installed life.

The initial hypothesis was that cycle-oriented utilization measures cumulated over the installed engine lifetime could be used to predict engine life. These measures were expected to be more powerful predictors of engine life than experience just before failure because engineering analyses suggest that engine stress is cumulative. The first of the two data files described in the previous section was used to test this hypothesis.

If the removal process is driven by the flying program, it should be possible to predict the probability of removal as a function of total hours accrued on an engine.[1] Because total hours is the sum of hours flown in the 30 days before removal and hours accrued before this period, a model of the form:

$$\text{Pr (removal)} = \beta_o + \beta_1 \text{ (recent hours)} + \beta_2 \text{ (past hours)} + \varepsilon$$

can be used to simultaneously test the effects of recent utilization and total flying hours.

---

[1]The purpose of this analysis was to establish whether utilization in the period immediately preceding engine removal influenced the probability of removal. The unit of observation was an individual engine during each month it was installed, with a categorical variable indicating whether the engine was removed during that month. Linear regression was used to test the relative effects of "recent" and "previous" experience. If "recent" hours appeared to have an effect, more rigorous analysis with other analytic techniques would have been indicated. In addition, the regression results show the effect of the monthly hour accumulation *rate*. If the removal process is driven by flying hours, the probability of removal should increase with increasing monthly flying activity.

Logistic regression is a maximum likelihood analytic technique that is frequently used to estimate the probability of an event as a function of several independent vaiiables. This technique constrains the value of the dependent variable to the range 0-1, and generates estimates of the coefficients of the independent variables that can be used to predict the probabi'ity of removal for an individual in the population.

An ordinary least squares regression model using a categorical indicator of whether or not failure occurred during a particular month as the dependent variable provides an approximation to the results that can be obtained using logistic regression. Such a model was specified for use with the data contained in the detailed (monthly) engine utilization data file described in the previous section. The results of fitting this model suggested that:[2]

- Recent flying hour utilization is not statistically significant as a predictor of engine removal probability;
- The coefficient for this variable is so small that it would make little difference even if it were significant;
- There is a slight negative relationship between previous hours and this probability, which probably reflects the effects of screening out short-lived engines in early months.

---

[2]The coefficients and t statistics for this model, as well as the other two mentioned in the text, are tabulated below:

Model	*Variable*	*Coefficient*	*t Statistic*
Flying Hour	Past Hours	-1.1127 (-05)	-4.27
	Recent Hours	-2.4873 (-06)	-0.04
Cycles	Past Cycles	-2.8221 (-05)	-3.82
	Recent Cycles	4.2675 (-03)	2.80
Both (Logs)	Past Hours	8.4363 (-03)	1.58
	Recent Hours	-6.8752 (-03)	-1.73
	Past Cycles	-1.4746 (-02)	-2.44
	Recent Cycles	2.1235 (-02)	4.31

The numbers in parentheses are the powers of ten needed to place the decimal point in the coefficient. The last model used logarithms of the parameters rather than their untransformed values.

In contrast, a model that used recent and previous cycles as the independent variables showed a statistically significant relationship between recent cycles and removal probability, as did a model containing both flying hours and cycles. This suggests that cycles are a better measure of engine stress than engine flying hours, and confirms the relationship expected for engineering reasons.

Since cyclic stress appears to be an important explanator of engine removals, a model that related the MTBR for an engine population to cycle-oriented measures of engine utilization could be used to predict the wartime removal rate. The preferred model would deal with individual engine removals, and would be of the form:

$$\text{MTBR} = \beta_0 + \beta_1 \text{ (cycle length)} + \varepsilon$$

Such a model, however, would be subject to specification error. Total hours is both the dependent variable and the numerator of the fraction to be used as the key independent variable, so positive correlation could be expected even if cycle length had no effect. Hence it is not appropriate to estimate a model of this form.

Although there are no good proxies for either total hours or cycle length that can be used to avoid this specification problem, a model using cycles per sortie and sortie utilization rate (in sorties/day) as independent variables to describe utilization was fitted to the data. This model, which explained only about 30 percent of the variation in engine life, suggests that increasing the average number of landings per sortie from one to four will reduce engine life by about 160 hours. It also suggests that increasing the sortie utilization rate can be expected to increase engine life.[3]

---

[3]The equation fit to the model was:

"Lifetime" hours =	720.8	- 164.2 (cycles/sortie)	+ 2757.5 (sorties/day)
Standard error	68.3	37.1	128.5
t Statistic	10.6	-4.4	21.5

Airline experience (see fn 2, Sec. IV) and other analyses (e.g., Berman et al., unpublished research) do suggest that increasing utilization decreases removals/flying hour. However, utlization ranged only from .004 to .69 sorties/day, and the bulk of the data was concentrated in

In view of the limitations of both the predictive power of this model and the data on which it was based, an alternative approach to the data analysis problem was pursued. The results of this analysis are described below.

A SURVIVAL ANALYSIS OF ENGINE REMOVAL DATA

The problem of estimating the MTBR of aircraft engines is quite similar to that faced by the actuary charged with developing life expectancy tables for insurance underwriting or the biostatistician comparing disease treatment alternatives. In fact, actuarial techniques have been used for the analysis of engine removal and failure data since World War II (Altman and Goor, 1948).

The actuarial analyst is concerned with the distribution of time to removal as well as the removal rate, which is simply the reciprocal of the MTBR for an exponential distribution of time to removal. Figure 13 shows the percentage of engines removed by 100 hr age interval for the data described in the previous section.[4] It suggests that the probability of removal decreases consistently after an initial increase during the early stages c engine life.

This decrease is expected, because engines that are removed during one interval do not have the opportunity to be removed during subsequent intervals. Although the probability of removal appears to be low for "old" engines, premature removals would truncate the distribution. Artificially constraining the distribution could affect the MTBR, hence the demand rate.

---

the range .05 to .24 sorties/day. It would be inappropriate to use this result to project to the expected wartime sortie rate. Furthermore, the result may reflect the fact that "good" airplanes are flown more often; causality may be the reverse of that implied by the model. In short, this model should *not* be used to project removal rates. The sole reason for describing it is to show that the effects of both variables, and in particular that used as a proxy for cycle density, are in the expected direction.

[4]Splines have been used to connect the data points in the plots shown in this section. In reality, only discrete step functions can be inferred from the data. Several methods for fairing a smooth curve through this step function are described in the literature (e.g., Gans, 1982; Miller and Singpurwalla, 1977; Parzen, 1962).

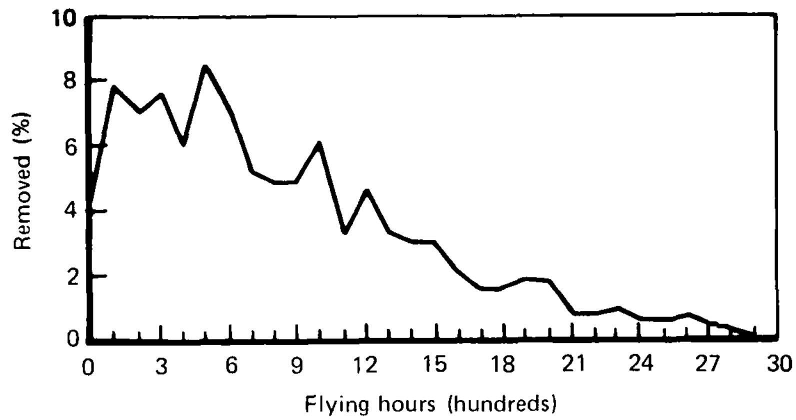


Fig. 13 — Percentage of engines removed by age interval

Figure 14 shows the cumulative percentage of engines removed by age interval. If the data are not censored, a "survival" curve can be constructed by plotting (100 - cumulative percentage removed) against age. The more complicated procedures needed to develop a survival curve for censored data are discussed in Sec. VI.

Although the MTBR for this sample is about 900 hours, nearly half of the engines are removed before they accrue 700 hours.[5] This reinforces the inference drawn from Fig. 13 that the tail of the

[5]The Air Force is currently using a wartime removal rate estimate of .91 engines per thousand engine flying hours. The implied MTBR is about 1100 hours, considerably higher than that observed in this dataset. A likely reason for this difference is that some engines were removed prematurely to be inducted for modification, but the removal cause codes in the engine history records attributed the removals to some other cause. This hypothesis is supported by a comparison of the removal reasons shown in the engine history records and those recorded by the GE technical representatives stationed at C-5 bases. The GE removal reason summary shows far more "scheduled" removals than the Air Force data. If some engines were removed prematurely, the data are censored, which would tend to reduce the computed MTBR. There is no way to determine for sure whether such censoring occurred, or which observations it may have affected. However, it is highly unlikely that the removal rate used in requirements computations *understates* representative peacetime experience. More probably, the MTBR developed from these data understates the true parameter value, because of unidentified censoring.

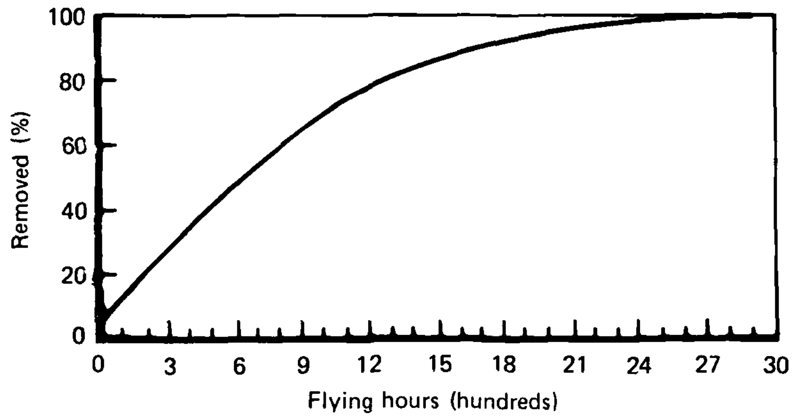


Fig. 14 — Cumulative percentage of engines removed

distribution contains much of the experience that is used to develop the demand rate.

Estimates of the conditional probability of removal within each age interval can also be developed from these data. This conditional probability, sometimes known as the "hazard function," is defined as the probability that a removal occurs during an interval, given that the engine entered that interval.

If the time between removals was exponentially distributed, a plot of the conditional probability of removal by age interval should trace out a horizontal line. Alternatively, if the hazard function is monotonically increasing (or decreasing), other distributional assumptions, such as the Weibull distribution, are appropriate (Hahn and Shapiro, 1967).

Figure 15 is a plot of the conditional probability of removal by age interval that corresponds to the data presented in Figs. 13 and 14. The line plotted between the points is a least squares linear regression of the conditional probability of removal on age interval. This line is not an estimate of the hazard function, but serves to emphasize that the "hazard" rises with increasing age. The increase in the hazard rate indicates that the exponential assumption implicitly applied in spares requirements computations is inaccurate.

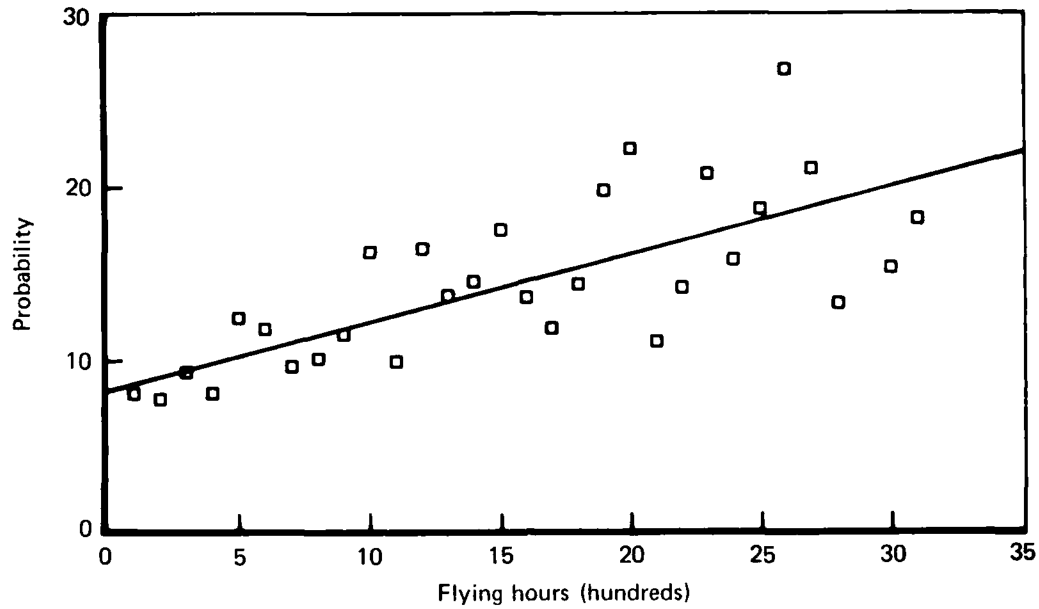


Fig. 15 — Conditional probability of removal by age interval

Even though analytic techniques such as those described above have been applied to engine data for years, little effort has been devoted to determining whether differences in utilization across engine populations result in significant differences in their hazard functions and survival curves. One such application could be to establish whether cycle length (or cycle density) affects engine life. If the hypothesis can be rejected that survival curves for two populations with different average cycle lengths come from the same distribution the techniques could also be used to predict the wartime MTBR.

The results of such an analysis are presented in Fig. 16, which is a plot of the survival curves for two subgroups of the engine population described in Figs. 13 through 15. The two groups were defined based on the average cycle length experienced over an engine lifetime on the assumption that "excess landings" (the number of landings in excess of the one required per sortie) generate one third of a full cycle.[6] The lower curve is based on a sample of about 500 engines with an average cycle length of less than 3 hours. The upper curve portrays survival probability for about the same number of engines that had average cycle lengths of more than 3.2 hours.

The difference between these curves appears to be small, but the results are statistically significant.[7] The total life expected from the engines in the "long cycle" group is more than 12 percent greater than in the other group--even though the difference in median cycle length is small.

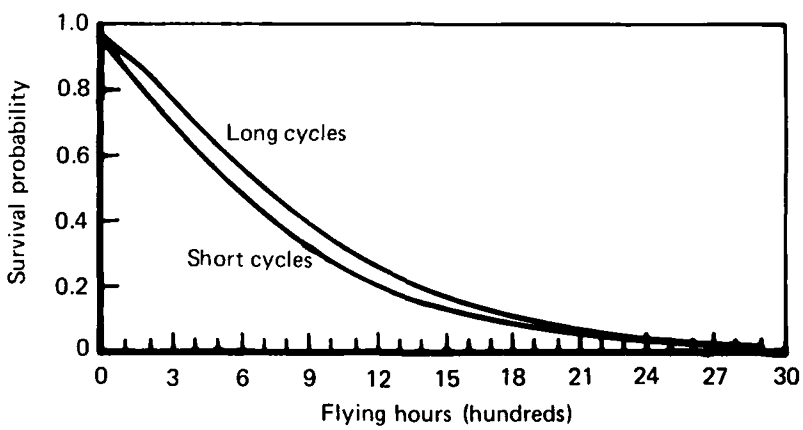


Fig. 16 — Survival curves for different cycle lengths

---

[6]This estimate may be conservative. "Partial" cycles that involve large throttle movements been estimated to cause half the stress of a full cycle (Byrd, Barrial, and Kostomay, 1978). Analyses of the the effects of F-100 engine throttle movements, which are generally smaller than those used during a landing, led to a conclusion that these changes generated one-fourth to one-third the effect of a full cycle.

[7]A Kolmogorov-Smirnov test (Siegel, 1956) supports rejecting the hypothesis that the two survival curves come from the same population at the .05 level.

The average life for the short and long-cycle engine groups is contrasted with that of the total population in Fig. 17. Projecting these results to the 4.5-5 hour average sortie (cycle) length expected in wartime yields an expected wartime MTBR of about 1050 hours, which is about 20 percent more than the average engine life realized in peacetime.

This result is consistent with the engineering discussion presented in Sec. III, and suggests that larger differences in average cycle length should result in correspondingly greater differences in survival probability. It is also consistent with airline experience.[8]

Hazard data can also be used to identify a theoretical distribution that fits the data. The increasing conditional probability of removal with age shown in Fig. 15 violates the exponential distributional assumption. The monotonically increasing hazard rate suggests a time dependency that is captured in the Weibull distribution. Several analysts have found that this distribution describes lifetime data for many complex items.

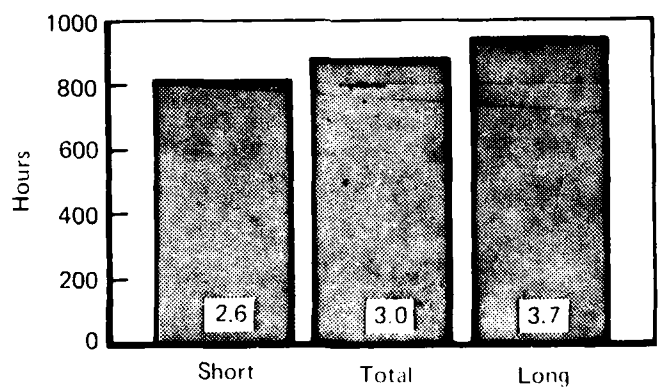


Fig. 17 — Mean time between removal for different cycle lengths

---

[8]Comparison of the data in Figs. 6 and 17 also reveals that Air Force removal rates are much higher than those experienced by the airlines. The Air Force has attributed this disparity to differences in military and commercial missions, but it is at least possible that Air Force maintenance practices also play an important role.

The Weibull is the limiting distribution for the minimum of *n* values from a large number of initial distributions bounded at the left (Hahn and Shapiro, 1967). Because a turbine engine is a complex item consisting of many components, and the engine fails when the first of these components fails, the Weibull distribution can be expected to fit empirical life data.

ESTIMATING THE PARAMETERS OF A WEIBULL DISTRIBUTION

The cumulative probability of removal for the Weibull distribution is given by:

$$F(t) = 1\text{-}\exp[(t - \gamma)/(\eta - \gamma)]^{\alpha},$$

where α is the shape parameter, η is the scale parameter or characteristic life, and γ is a location parameter. The failure rate is a power function of time for this distribution.

Several graphical methods have been developed to facilitate visual tests of the fit between empirical and theoretical data. These methods take advantage of the fact that $\ln\{1/[1 - F(t)]\}$ will plot as a straight line against ln(time) if the distribution is Weibull (L. S. Nelson, 1967). The location parameter accounts for a period during which no failures occur. It is usually assumed to be zero, unless there are physical reasons to believe there is a minimum life, or successive trials indicate that adjusting failure time by a fixed amount "straightens" the plot.

One of these methods is to plot the cumulative percent failure against time on a Weibull probability chart.[9] The shape and scale parameters for the distribution can then be estimated directly.

---

[9] Another method is to plot the cumulative hazard number, defined as the sum of the conditional probabilities of removal to time t, against $\ln t$ on special graph papers. If the data fit a Weibull distribution, the hazard number should plot as a straight line against t on Weibull probability paper (Nelson, 1969). A third approach requires plotting the total time on test vs. t and using templates to fit a theoretical distribution to the resulting curve (Barlow and Campo, 1975). The results of applying this technique are equivalent to those of the other methods, but the graphical evidence of the fit between the data and the hypothesized distribution is less readily apparent.

Figure 18 presents such a Weibull probability plot for the engine failure data. It indicates that the data fit a Weibull distribution with a scale parameter of 960 and a shape parameter of 1.33, and the survival function for these data can be approximated by:[10]

$$S(t) = \exp[(-t/960)^{1.33}]$$

This function is plotted in Fig. 19. The survival curve from the engine failure data is also plotted, but it is difficult to discriminate between the two curves.

Survival curves developed from this peacetime operational experience cannot be used to estimate the wartime TF39 removal rate because:

- Data at expected wartime cycle lengths are not available
- The engine has undergone a major modification that is expected to increase the MTBR since these data were collected.

These analytic techniques can be used to both design and interpret the results of the operational tests that are needed to develop a removal rate estimate.

[10]Grouped data were used to develop the Weibull probability plot. A maximum likelihood procedure developed by William Rogers of the Rand Corporation (Rogers and Hanley, 1982) was used to estimate the parameters from the individual data points. This procedure yielded estimates for the shape and scale parameters of 1.36 and 982 respectively. This procedure was also used to estimate the parameters of the distributions plotted in Fig. 16. For the short cycle group, the procedure yielded estimates of $\alpha = 1.27$ and $\eta = 914$, which are close to the values of 1.26 and 890 estimated from the grouped data. The maximum likelihood estimates for the long cycle group were $\alpha = 1.43$, $\eta = 1046$ vs. 1.39 and 1040 with the graphical technique.

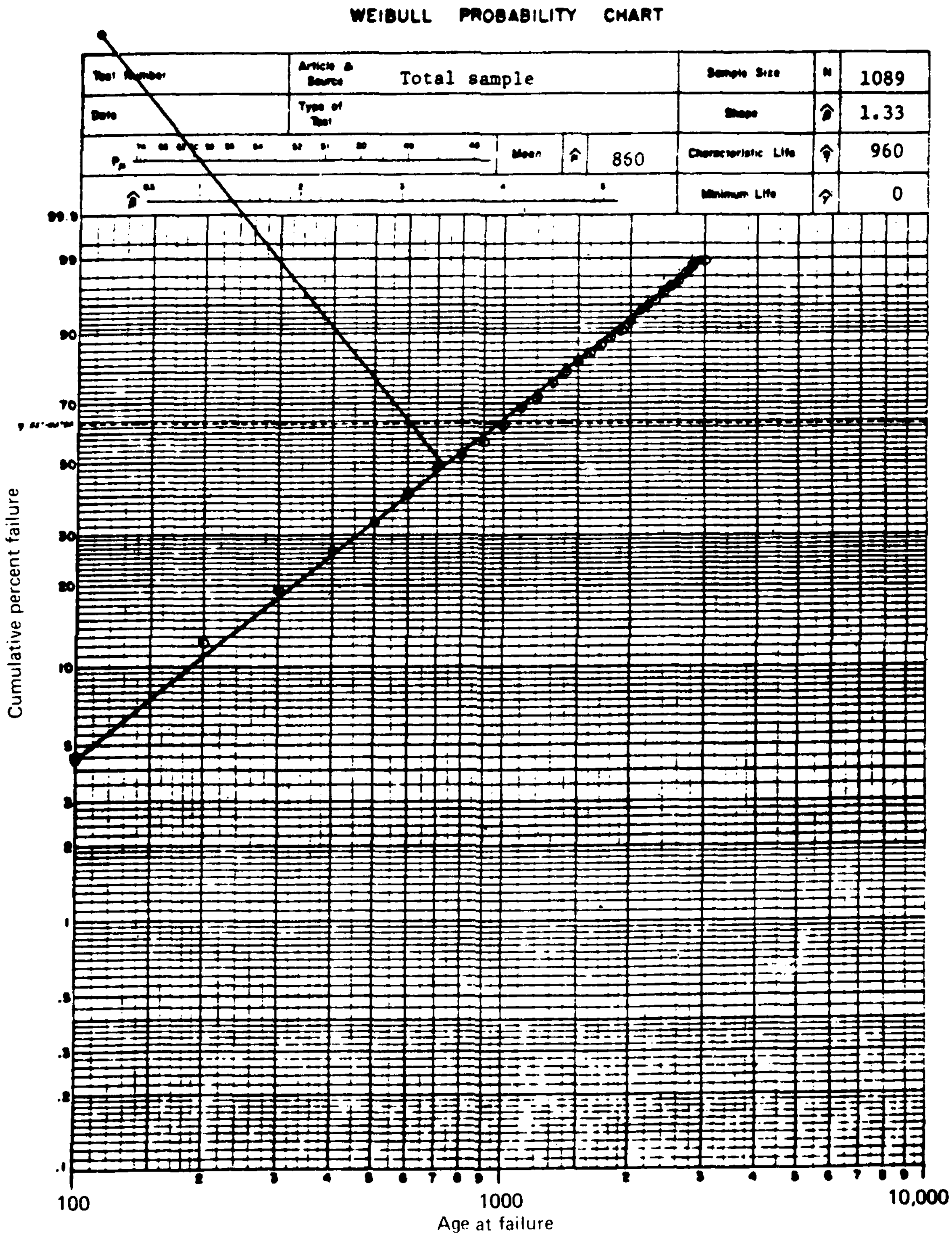


Fig. 18 – Weibull probability plot

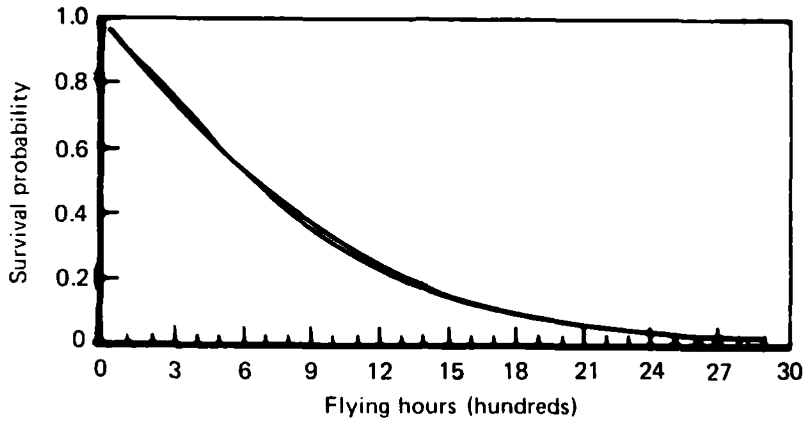


Fig. 19 — Predicted and observed survival curves

VI. GENERATING DATA FOR LOGISTICS PARAMETER ESTIMATION

Despite a considerable body of evidence that increasing cycle lengths should reduce engine failure rates, requirements computations continue to be based on the assumption that removals are driven by the flying program. This may be partially attributable to the fact that other models are not available. Although the GE engineering model could be used to project the wartime failure rate, military managers would probably be reluctant to base their requirements estimates on the predictions of an engineering model without some empirical validation of its predictions.[1]

The model's predictions could be tested, or the wartime removal rate could be estimated directly, by means of an experiment designed to accumulate data concerning the effects of cycle length on observed engine life. Such an experiment could be conducted without increasing the level of C-5 flying activity by changing the scheduling rules used to assign aircraft to missions for the approximately 50,000 hours flown each year. Thus no additional flying would be needed to test the cycle length hypothesis, although changes in the policies used to assign aircraft to missions in peacetime would be necessary.

This section describes an experiment that could be used to test the effects of varying cycle lengths. It also outlines a Bayesian approach for estimating the wartime removal rate and discusses the implications of the experimental design for aircraft scheduling policies.

AN APPROACH TO THE EXPERIMENTAL DESIGN PROBLEM

The problem of designing an experiment to explore the effects of utilization on aircraft engine life is quite similar to that of designing a clinical trial to test the effects of treatment alternatives on medical patients. Such "life tests" require sample sizes large enough to permit statistically valid conclusions to be drawn from

[1]The model has been validated empirically, as was discussed in Sec. III, but the Air Force would probably require validation in a military setting.

experimental results. In addition, they take a considerable amount of time to accumulate sufficient data. Medical studies usually follow their patients for several years to establish the relative effectiveness of treatment alternatives.

Typically clinical trials in medicine are structured to support direct comparison of treatment alternatives; researchers usually do not have a basis for estimating the survival curves for the alternatives before the experiment begins. One approach to the problem of estimating engine life as a function of "treatment" would be to design a test similar to those used to discriminate among treatments in medicine. The results of the test could be used to project the wartime removal rate.

The same data could be used to test the GE model's predictions. An engine survival curve could be developed from the wartime removal rate projection and compared with actual test results. The preliminary estimate could then be updated as experience accumulated with Bayesian methods. This approach would enable the Air Force to develop wartime removal rate estimates before the test was complete.

Any experimental design involves decisions concerning:

- The number of subjects (engines) to be used in the experiment;
- The method to be used to assign subjects to treatments;
- The length of the test;
- Techniques for extracting the maximum amount of information from the data before all of the results are in;
- Procedures to avoid introducing bias into the experiment;
- The analytic techniques to be applied to the experimental results.

The number of engines needed for the test is a statistical issue and should be based on the degree of statistical validity desired. The sample needs to be large enough to permit observation of sufficient events, where the events of interest for survival analysis are "deaths" rather than trials (Freedman, 1982). Thus the sample size needs to be adequate to provide: (1) statistically valid results when the desired number of deaths have occurred and (2) a reasonable probability that this number of deaths will have been observed at the time selected for analysis.

Most experimental designs that are intended to compare two or more treatments require a randomization protocol to ensure that test results are not biased (Armitage, 1982). Bias could be introduced if there are differences in the characteristics of subjects that have nothing to do with the variables being tested. So long as none of the engines selected for test contain components with life limits that might require them to be removed prematurely, there should be no such differences among engines produced by the depot or a particular JEIM shop. Minimum limits on the remaining life of components used in engine repair are established by policy (Department of the Air Force, 1982). Hence selecting an engine from stock or JEIM production to meet an aircraft requirement constitutes random assignment; a formal randomization scheme should not be required for this test.

The calendar time needed for the test will be determined by: (1) the length of time needed to bring the required number of engines onto test and (2) the number of failures to be observed to yield the desired levels of statistical power. It would take over a year to introduce a large sample of engines to test,[2] and years longer for all of the engines used in a treatment experiment to fail. Fortunately it is not necessary to wait that long to perform the analysis. Interim results can be obtained using the Bayesian methods described in this section and illustrated in Appendix D. Reasonably complete results for classical statistical analysis will be available after only part--preferably 60 to 70 percent--of the engines have been removed.

Terminating the experiment (for analysis purposes) while 30 percent of the engines remain in service imposes Type II censoring on the data. However, the product-limit (PL) estimate of survival probability has been shown to be the maximum likelihood nonparametric estimate for this probability even when the data are censored. The PL estimate exploits the fact that units either failing or removed from trial during one interval are not susceptible to failure during subsequent intervals.

---

[2]The test could be initiated more rapidly if several engines were refurbished and installed before removals were necessary. This would create an unreasonable burden on maintenance.

The product-limit estimate of the probability of survival to time *t* is:

$$P(t) = \prod_r [(N - r)/(N - r + 1)]$$

where r includes the positive integers for which tr' < t, and tr' is the age at failure. The expression allows for the possibility of censored observations (Kaplan and Meier, 1958). This method of treating censored data will be assumed for all of the discussion that follows.

A "CLINICAL TRIAL" OF ENGINE TREATMENT ALTERNATIVES

If training sorties cause greater engine stress than cargo sorties that have longer average cycle lengths, the survival curves for engine populations exposed to the two operating environments should differ. The data presented Sec. V suggest that survival probabilities for engines treated with long cycle lengths have higher survival probabilities, particularly at low age, than those cycled more intensively.

One approach to establishing the effect of cycle length on engine life would be to assign *n* aircraft to cargo missions and *m* aircraft to training missions for the duration of the experiment. Experimental aircraft would be assigned to fly most cargo missions, and the balance of programmed C-5 flying hours would be distributed across the remainder of the fleet. The number of aircraft required in each group would be determined by the size of the sample needed to distinguish between the two survival curves.

The logrank test is frequently used to test the difference between two life distributions and is particularly good when the ratio of the hazard functions for the two treatment groups is approximately constant over time (Peto et al., 1975, 1976).[3] The test statistic is essentially a comparison of the observed number of failures in each group and the "expected" number assuming no difference in the two curves (Kalbfleish and Prentice, 1980). Minimum sample sizes needed to reject the hypothesis that two survival curves come from the same distribution using the logrank test have been tabulated (Freedman, 1982). An extract from Freedman's table is reproduced as Table 1, which shows the sample size (number of engines) needed to reject the hypothesis that the "control" and "experimental" survival curves come from the same population at various levels of confidence and test "power."[4] These sample size requirements assume that half of the number shown in any given cell of the table are designated "controls," and the other half are included in the experimental group. Normal theory can be used to compute the effective experimental sample size when the populations in the experimental and control groups differ in size using the expression:

$$2(1/n^{*}) = (1/n) + (1/m)$$

To apply the table the analyst must specify:

- The control group survival probability at which the analysis is to be conducted,
- The expected difference in survival probability between the two groups when the appropriate percentage of the control group has "died."

The experiment should continue until at least 60-70 percent of the control engines have failed. The calendar time needed to attrit 70 percent of the controls is a function of: (1) the utilization rate of the test aircraft; and (2) the distribution of time to removal for the TF39-1C engine.

---

[3]The test can also be used when the two groups do not face "proportional hazards." The differences between the shape and scale parameters estimated for the distributions fit to data from the "short" and "long" cycle groups described in Sec. V indicate that the hazard functions for the short and long cycle groups are not proportional. Engines exposed to short cycles have a shorter characteristic life, but they also have a smaller shape parameter. Thus the hazard functions for the two groups should converge beyond the characteristic life of the long-cycle group.

[4]Tabulated values refer to the number of engines; one-fourth the number of aircraft would be needed. In parentheses are the expected number of failures experienced across the total sample for that cell of the matrix.

Table 1

SAMPLE SIZE REQUIREMENTS FOR THE LOGRANK TEST

Sample size needed to detect an improvement (P_2 - P_1) in survival rate over a baseline survival rate (P_1) with a one-tailed test when:

i. $\alpha = .05,\ 1 - \beta = .80$
ii. $\alpha = .05,\ 1 - \beta = .90$
iii. $\alpha = .01,\ 1 - \beta = .95$

	$P_2 - P_1$				
P_1	0.05	0.10	0.15	0.20	0.25
	1439 (1115)	398 (298)	194 (140)	118 (83)	82 (55)
0.20	1993 (1545)	551 (413)	268 (194)	163 (114)	113 (76)
	3668 (2843)	1013 (760)	492 (357)	200 (210)	207 (140)
	1723 (1249)	464 (325)	221 (149)	133 (86)	96 (56)
0.25	2386 (1730)	643 (450)	306 (206)	183 (119)	133 (78)
	4391 (3183)	1132 (827)	562 (379)	336 (218)	244 (142)
	1960 (1323)	518 (336)	242 (151)	143 (86)	96 (55)
0.30	2714 (1832)	717 (466)	335 (209)	198 (119)	133 (76)
	4995 (3371)	1319 (857)	616 (385)	364 (218)	244 (140)
	2147 (1341)	559 (335)	258 (149)	151 (83)	100 (52)
0.35	2973 (1858)	774 (464)	357 (205)	208 (114)	138 (72)
	5471 (3419)	1423 (853)	656 (377)	82 (210)	253 (133)
	2281 (1312)	586 (322)	268 (140)	155 (77)	102 (48)
0.40	3159 (1816)	812 (446)	370 (194)	214 (107)	140 (66)
	5814 (3343)	1493 (821)	680 (357)	392 (196)	257 (122)
	2363 (1240)	600 (300)	271 (129)	155 (70)	101 (43)
0.45	3272 (1718)	831 (415)	375 (178)	214 (96)	140 (66)
	6022 (3161)	1529 (764)	689 (327)	393 (177)	255 (108)

Source: Freedman, 1982.

There are two strong arguments for using normal peacetime training experience as tle control for the proposed experiment:[5]

- Only part of the test subjects would need to be identified explicitly,
- The calendar time needed for the test could be reduced relative to that required if two groups of engines were placed on test at the same time.

The average utilization rate of the C-5 fleet has been less than three hours per aircraft per day for the past several years, and reliability improvement was part of the justification for the modification that converted the TF39 to the 1C configuration. If the MTBR for the modified engine is in the 1400 hour range,[6] as is expected, it would take nearly two years to generate 70 percent of the removals from the control group at recent utilization rates.

Clearly the time needed to accumulate data on control engines could create a testing constraint. However, life data for the TF39-1C engines that have been previously or are currently installed in the fleet can be used to construct a control group survival curve. This curve could provide a basis for comparison with that for test engines as failures occur. Using this experience for an experimental control could also increase the effective size of the experimental sample.

Estimating the expected improvement in experimental group survival probability when 70 percent of the control engines have been removed is a more difficult problem. Two sources of data relevant to this problem are available:

---

[5]The data described here cannot be used to construct the control survival curve because they predate the TF39-1C modification. Using peacetime (TF39-1C) experience rather than training flights as the control would tend to reduce the expected difference in survival probability between the control and experimental groups, but this cost is more than offset by the benefits outlined above.

[6]The analysis in Appendix A demonstrates that no additional spare engine procurement is necessary if the expected MTBR for the TF39-1C engine is realized.

- The airline experience described in Sec. III,
- The analyses presented in Sec. V.

A linear interpolation of the airline data indicates that the expected MTBR for engines exposed to 4.7 hour sorties is more than 30 percent greater than that for engines used on 2.9 hour sorties. A similar projection of the C-5 engine life data analysis results suggests that there should be at least a 20 percent difference in MTBR for the control and experimental average cycle lengths.[7]

These facts, coupled with the statistical results indicating that cycles are a better measure of engine stress than flying hours and that increasing sortie utilization rates increases engine life, suggest that it is not unreasonable to expect a .15 difference in survival probability when 30 percent of the control engines survive.[8]

Assuming that 18 aircraft can be designated as test platforms (n = 72) and that engines installed on other aircraft, as well as previous TF39-1C removals, are treated as controls (m = 436), the effective total sample size ($2n^*$) is 247. Table 1 indicates that this sample is sufficient to reject the null hypothesis with 80 percent power when: (a) 70 percent of the control engines have been removed and (b) 45 percent of the experimental group survive.

---

[7]Both of these estimates are based on fitting a line to the MTBR data using median cycle length to represent group experience. Depending on the method used to introduce engines to test and the number of pre-test failures used to increase the effective sample size, the median control cycle length should be between 2.4 and 2.7 hours per cycle. The experiment should be structured to ensure that the median cycle length for experimental engines is about 5 hours.

[8]If half of the available C-5 flying hours are flown by the experimental group with average sortie (cycle) lengths of 5 hours, and the remaining aircraft make the balance of the landings that are made in a typical year, the average cycle length for the control group will be about 1.9 hours. Averaging this value with that for an assumed 200 prior failures with average cycle lengths of 3.1 hours (the average observed in the data), the overall average control group cycle length will be about 2.45 hours. Extrapolating the difference between short and long cycle survival probability to the cycle length expected for the experimental group yields an expected difference in survival probability of nearly .16 when about 30 percent of the controls survive. This calculation does not account for any possible increase in experimental group survival probability attributable to an increase in aircraft sortie rate, suggested by the regression results described in Sec. V.

Discrimination between the two survival curves could be improved by increasing the size of the experimental group. However, it is unlikely that more aircraft could be assigned as test platforms without creating an unacceptable constraint on peacetime aircraft scheduling. This need not limit the experiment even if the difference between experimental and control group survival probability is less than .15 at the analysis point. Although the tabulated values for the logrank test were used to address the sample size issue, other tests can be used to analyze the experimental results.

In particular, the Kolmogorov-Smirnov test could be used in the analysis. This test is the most powerful for detecting many possible differences between two distributions (Siegel, 1956) and is more sensitive to differences in survival curves than the logrank test. It was not used to address the sample size issue because it is easier to infer sample size requirements for the logrank test.

Minimizing the number of engines that must be monitored during an experiment would limit its effect on peacetime aircraft scheduling.[9] It should also reduce the potential for operator-induced bias in the experimental results. The risk of bias could be further reduced if new scheduling rules were promulgated without disclosing the purpose of the test.

The experiment would be conducted by assigning "test" aircraft to fly cargo missions exclusively, with minimum average sortie (cycle) lengths of four hours. Although it would take a great deal of time for the test aircraft to accumulate experience at recent aircraft utilization rates, utilization has been constrained because of wing cracking problems that are in the process of being corrected through a wing modification program. Thus it should be possible to accelerate the test by allocating more of the available C-5 flying hours to the test aircraft.

---

[9] As has already been noted, aircraft schedulers have done a remarkable job of equalizing the distribution of missions across aircraft because of the C-5 wing cracking problem. Rules designed to induce variation in utilization should present less of a scheduling problem, particularly if only a few aircraft are affected.

In peacetime aircraft are assigned to a home station at either Dover or Travis Air Force Base. If half of the test aircraft were assigned at each of these bases, average utilization rates of nearly four hours per day on normal channel missions should be possible.[10] The fact that these aircraft would not be available for training missions would create some scheduling constraints, but these constraints should be less binding than those aircraft schedulers have had to confront for the past several years.

A final factor that must be considered to estimate the calendar time needed for the test is the procedure to be used to introduce engines to test conditions. It would be unreasonable (if not impossible) to refurbish and change over 70 engines in a short period of time simply to begin an experiment, regardless of its importance. However, waiting for unscheduled removals to occur could take a very long time just to introduce the required number of engines to test even at an increased aircraft utilization rate. Fortunately an alternative is available. The wing modification program is producing an average of 1.5 aircraft per month. Each of these aircraft must be fitted with four engines before it is returned to the Air Force. The required number of engines could be brought on test by installing repaired engines from stock on the aircraft generated by the modification program. Some of these engines might come from aircraft introduced to test earlier after opportunistic engine replacements. This would permit all of the necessary engines to be placed on test within a year and would have the additional advantage of providing a "lead the fleet" test of the wing modification itself.[11] Preliminary results of the test would be available within a matter of months, and fairly complete results could be expected within two years.

---

[10]The Military Airlift Command (MAC) routinely serves several overseas channels in peacetime. For example, several C-5 flights a week are scheduled from Travis to Clark AFB in the Phillipines, and several more fly from Dover to Europe. Two or more sorties are flown on each of these missions.

[11]The first aircraft to receive the wing modification is already being used in a "lead the fleet" program.

This proposal for a test addresses the experimental design issues listed earlier. To recap, two changes in current aircraft assignment and utilization policies are needed to structure a "clinical trial" of the effect of cycle length on engine removal rate:

- Assign 18 aircraft to fly cargo missions exclusively,
- Increase utilization rates for these aircraft.

Some data for the control group would be available when the test was initiated, so results for the test group could be compared with control group data during the course of the experiment with any of several available tests. The Kolmogorov-Smirnov test is a simple procedure for testing the hypothesis that two survival curves have been generated from the same distribution. Procedures for developing nonparametric confidence limits for the survival curve have also been developed (Kaplan and Meier, 1958; Simon and Lee, 1982). The log-effect function, the ratio (in logarithms) of the hazard rates faced by the two groups of engines, is a simple test for proportional hazards (Kay and Schumacher, 1983). It should also reveal divergence between the two hazard functions. Finally, estimates of the log survivor and hazard functions could be made directly (Rice and Rosenblatt, 1976).

TESTING AND UPDATING PREDICTIONS OF THE GE OPSEV MODEL

GE's OPSEV model could be used to predict the wartime removal rate. This prediction, coupled with an assumption concerning the distribution of time to removal, could be used to construct a survival curve that could be tested in an experiment.

For example, assume that the model predicted a removal rate of .7 engines per thousand engine flying hours, and that engine life is believed to fit a Weibull distribution with a shape parameter of 1.3 and a scale parameter of 1550. The resulting "predicted" survival curve is shown in Fig. 20. An exponential distribution with parameter λ=.0007 is also shown to contrast the classic exponential assumption with the Weibull survival curve. The apparently minor differences between the two curves suggest that the exponential assumption is probably adequate for most purposes. However, *life data analyses that use parametric* methods should use the distribution that best fits the data. The Weibull distribution appears to meet this criterion for engine life data.

This procedure would permit comparison of experimental results with predictions made using the engineering model as well as the comparisons of control and test group experience. Actual failure data accumulated during the test could be used to construct an empirical survival curve using the methods described in Sec. V. This curve could then be compared with the estimated curve using the statistical tests mentioned above.

This procedure could be used initially to validate the model's predictions and ultimately to develop a new removal rate projection. An alternative would be to use Bayesian methods to update estimates of the parameters of the Weibull distribution during the course of the test. The methods provide a means for continuing update of the removal rate estimate as removals occur.

Bayesian methods take advantage of the information available before the test begins. Furthermore, they can be used to estimate the value of the information obtained by test (Raiffa, 1968). A great deal of work has been devoted to applying these methods to reliability problems (Shimi and Tsokos, 1977), and the reliability methods have addressed

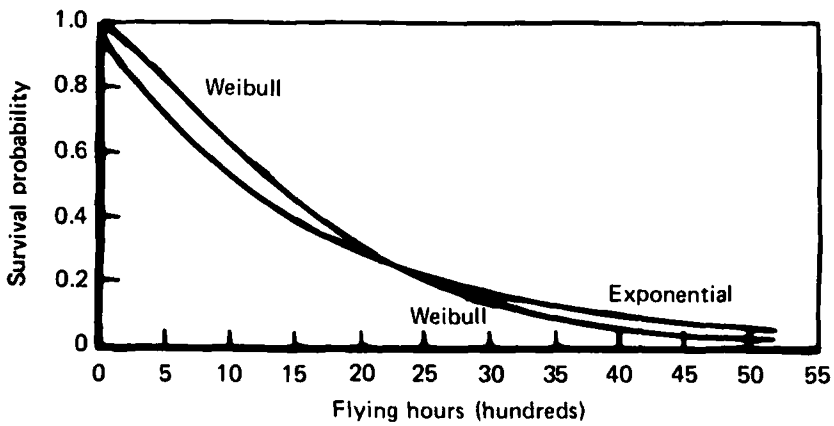


Fig. 20 — Exponential and Weibull survival curves

procedures for applying them to the Weibull distribution (e.g., Mann, 1970; Soland, 1968, 1969). Appendix D provides an example of the calculations required.

VII. FORECASTING LOGISTICS RESOURCE REQUIREMENTS: POLICY IMPLICATIONS

The effort to address the C-5 engine requirements problem has exposed two larger problems that must be addressed by policy:

- The "statistical" problem of estimating the wartime demand rate implicitly requires a policy decision,
- The basis for estimating this parameter can be improved by changing aircraft scheduling policies.

Policy decisions (including those couched as assumptions) concerning statistical issues are irrational if there are better alternatives available. Although there will always be some uncertainty concerning the validity of statistical models, and the appropriate model is likely to vary among components, the resource implications of the wartime resource requirements estimation problem demand improvements in the analytic basis for requirements estimates.

This need is not confined to the C-5 engine requirements problem. Requirements for other resources are probably inaccurate for reasons analogous to those that affect the engine requirements issue. The approach outlined here should be applied to other resources.

The major difference between peacetime and planned wartime utilization of the C-5 aircraft makes estimating C-5 spare engine requirements an important issue. Although parallel problems in estimating other support resource requirements may not be as dramatic, the problems, potential solution approaches, and sources of resistance to change parallel those addressed in the C-5 context. A framework for addressing these problems can be summarized as follows:

- Component removal rates observed in peacetime reflect the combined result of several individual component failure modes,

- At least some of these failure modes can be associated with aircraft (and component) utilization,
- Utilization patterns in wartime will differ from peacetime experience,
- Peacetime simulation of wartime operating conditions can provide a basis for:
  - Estimating the relationship between utilization parameters and component life,
  - Developing logistics parameter estimates for use in wartime resource requirements computations.

AIRLIFT ANALOGIES TO THE C-5 ENGINE REQUIREMENTS PROBLEM

Utilization rates for all of the Air Force's aircraft are expected to increase in wartime. The rate of increase is particularly important for strategic airlift aircraft because peacetime training requirements are low relative to the airlift capacity required in time of war.[1] Requirements for components other than engines used on these aircraft should also be affected by the major change in utilization and landing rates that will occur in wartime.

Structural components, especially landing gear, should reflect particularly noticeable effects, even though wartime landings will involve heavier loads. However, the resource implications of changing base-level removal rates for landing gear components will be small because most landing gear repairs are accomplished during Periodic Depot Maintenance (PDM) in peacetime (Embry, 1982). These components are also much less expensive than engines and engine components.

Changes in landing-related stress could also affect avionics components, which account for a large fraction of both peacetime and projected wartime maintenance requirements. The observed increase in avionics maintenance requirements associated with frequent landings has

[1]The peacetime flying program is structured to meet minimum pilot proficiency needs. The airlift fleet has a great deal of transportation capability that could also be used in peacetime, but its use is limited to avoid placing the service in direct competition with private industry.

prompted the Navy to (unofficially) designate specific aircraft for multiple-landing missions ("bounce" aircraft). The avionics systems on these aircraft are permitted to degrade during peacetime training operations.

Avionics components also account for a large fraction of computed war reserve investment needs for strategic airlift aircraft. If demand rates for these components can be expected to change in wartime for reasons analogous to the cycle density effect on engines, revising parameter estimates would also greatly affect avionics war reserve requirements.

A reduction in the number of landings per flying hour is not the only reason to believe that the demand for avionics components, at least for cargo and bomber aircraft, would decrease in wartime. Although the reasons for failure of these components may not be as well understood as they are for engines, the demand for electronic items has been observed to decrease with increasing utilization in a wide range of applications. Shurman's time-dependent model of aircraft maintenance requirements was developed primarily from avionics data (Shurman, 1979); Shaw's estimates of the demand-sortie length relationship applied to avionics as well as engines (Shaw, 1982); airline experience suggests that maintenance manhours per flying hour decrease with increasing utilization (Carlyle, 1976); maintenance manhour per flying hour analyses in military settings confirm this experience (Donaldson and Sweetland, 1968); and the demand for shipboard electronic components on vessels deployed off the coast of Viet Nam was lower than that observed during routine peacetime steaming.[2]

One reason for such changes in the demand rate for avionics items might be that components are subjected to fewer power surges per flying hour (power surge density is lower) when utilization rates increase. In fact, the basic premise of accelerated life testing is that stress affects component life, and testing under increased levels of stress can provide information concerning life distributions under normal levels of stress.[3]

---

[2]Conversation with LCDR James McClure, Supply Officer of a surface combatant during the Viet Nam conflict, summer 1974.

[3]Accelerated life tests are sometimes used to explore component failure modes. Methods to use the results of these tests to predict service life have also been proposed (Barlow and Schener, 1971; Kahn, 1977; Singpurwalla and Al-Khayyal, 1977).

Engineers undoubtedly already know a great deal about the effects of utilization on the failure modes of avionics items, but there was no opportunity to explore the existing state of knowledge during the course of this research. It is an area that warrants investigation similar to that proposed for C-5 engines. In fact, the tests proposed for establishing the wartime demand rate for engines could also generate data that could be used for research into the effects of utilization on the life of other components installed on strategic airlift aircraft.[4]

ANALOGIES IN OTHER SETTINGS

If utilization affects the demand for components used on strategic airlift aircraft, it should also affect that recorded for tactical aircraft. In some instances, even though tactical training syllabi have been designed to approximate wartime requirements, wartime demand rates may increase over those experienced in peacetime. Tactical pilots may find it necessary to increase throttle use in response to combat situations, which could result in an increase in the wartime engine demand rate. Equipments might also be subjected to greater G forces than they are in peacetime.

Increasing utilization could also decrease demands for some components by limiting the effects of changes in the equipment's operating environment. Increasing the sortie rate for an individual aircraft could reduce the effects of temperature changes on some equipments,[5] and limit the frequency of power transients experienced by others. Operational schedules that required use of all aircraft would limit the opportunity for hydraulic seals to dry out, etc., as sometimes happens in peacetime when the operating unit can afford to use only some of its aircraft.

---

[4]Survival analysis similar to that proposed for engines would be appropriate. Some of the early work dealing with electronics reliability characteristics found that the Weibull distribution fit life distributions for electron tubes (Kao, 1959), and similar results were obtained for transistor technology. Modern integrated circuits are not subject to the same failure modes as earlier electronics circuit designs, but utilization probably affects the life of these components as well, and theoretical distributions could be fitted to their survival data.

[5]Because temperature affects avionics life (Jones, 1983), temperature changes will probably also have an effect.

Although research is continuing, the changes in malfunction rates observed in two F-16 squadrons during a surge exercise have yet to be explained (Embry and Crawford, 1983). Current research is investigating a "good vs. bad airplane" hypothesis;[6] additional effort might be expended profitably to improve understanding of equipment failure modes and their relationship to utilization. The possible effects of such changes could be explored using means such as those proposed for testing the effects of utilization on C-5 engine requirements. There is no reason to believe that the "linearity" assumption has any more validity in the tactical than the airlift context.

OTHER APPLICATIONS OF LIFE TEST RESULTS

The discussion here has emphasized the implications of wartime demand estimates for the problem of developing war reserve inventory requirements. In reality, these parameter estimates have more far-reaching implications. Requirements for test equipment and maintenance manpower in the field are based on forecasts of maintenance workload that depend on this same parameter. Depot maintenance capacity requirements are also based on wartime workload projections. Thus a change in demand rate estimates would influence not only spares but manpower, facilities, and equipment requirements. There are no data available to estimate the amount of money spent on support resources whose "requirements" are based on wartime demand rate estimates, but it is safe to say that total costs are measured in billions of dollars.

Changing demand rate estimates would not only influence the costs of provisioning a system to support the operating forces with the current support structure but would change perceptions of the viability of support structure options. Previous analyses have shown that wartime

---

[6] "Good" airplanes are those that experience few failures. A natural selection process operates to ensure that these good airplanes fly more often than those with higher failure rates. One possible explanation for the decrease in removal rates observed when activity levels increase is that the "bad" airplanes are removed from the sample early in the exercise. Even if this hypothesis is verified, it does not address the question of *why* some aircraft have higher failure rates than others.

capability could be improved and support costs reduced if such structural options were implemented (Drezner, 1975; Shulman, 1980).

ENGINEERING PROJECTIONS VS. OPERATIONAL EXPERIENCE

Engineering models and reliability predictions are used routinely during the design process. One purpose of acceptance testing is to ensure that MTBF goals can be met (Department of Defense, 1977). Thus the services already have experience in using the results of tests (experiments) to validate the predictions of engineering models.

Except in the case of reliability improvement warranties and modification programs, however, little use is made of engineering projections after the government has accepted an equipment. In-service life distribution data are considered only through the indirect means of developing demand rate estimates once a system is operational in the field.

This Note provides evidence that continued analysis of life data could improve the basis for estimating peacetime as well as wartime logistics parameters, and that engineering estimates can play a useful role in these analyses--particularly if aircraft scheduling policies are modified to reflect the need to generate meaningful logistics data.

Continued testing is particularly important because original acceptance tests do not predict reliability in the field very well. On one hand, acceptance test conditions seldom approximate field operating conditions (Cote and Birkler, 1979). On the other, complex items usually undergo modification programs that enhance their reliability. The "hazard curve" tends to decline toward the horizontal curve implied by the exponential distribution as a result of these modification programs (Nowlan and Heap, 1976).

Continued attention to life distribution data could hasten identification of the failure modes that are responsible for poor field reliability. This could aid the redesign process that is implemented through modifications and support "maturational development" of military equipment.[7]

---

[7]The term "maturational development" was coined in a proposal for improving field performance by sequential development of subsystems independent of platform development. This approach to design is taken with engines and could profitably be applied to avionics system

The testing program proposed here is really no more than an expanded acceptance test procedure that would use engineering projections in Bayesian analyses of acceptance trial data. Coppola (1981) proposed such an approach for initial acceptance testing.

INCENTIVES FOR CHANGE

Demands for considering support requirements early in the acquisition process and for making "supportability" a design requirement have increased over the past few years (e.g., Mullins, 1983; Rice, 1979; Rich and Drezner, 1982; Walker, 1980). This interest is attributable to the rising costs of support as well as to the recognition that support capability may be the Achilles heel of combat capability in future conflicts. The need for increased emphasis on support, however, is not confined to the acquisition process. Operational practices should be designed to accumulate information relevant to the resource requirements problem. This should entail modifications in operating policies to maximize the amount of information that can be obtained from normal peacetime operating data.

The algorithms used to schedule aircraft in peacetime are usually designed to achieve operational (training) objectives, subject to the constraints created by material problems or maintenance availability. Thus logistics does create constraints, but scheduling operations to support collection of logistics data is not one of them. In fact, "surge" exercises are typically preceded by a massive effort to accumulate extra support resources to ensure that logistics problems do not constrain operational flying activity. There is even some evidence that equipment failures may not be reported when they might require that the aircraft be removed from the flying schedule (Lippiatt, 1978).

---

development (Shulman et al., 1981). The proposal is to extend the approach to equipment already in the field and use life data analysis techniques to quantify the value of design improvements. The redesign problem also presents an opportunity for combining engineering projections and operational data to improve design (Jones, 1983) and logistics parameter estimation.

Data collected to improve requirements estimation could reduce the level of assets for support operations, which is a disincentive to explore the relationship between flying activity and maintenance demands. Operators are not forced to confront the tradeoff between investments in support and in new weapons procurement. In fact, they do not even incur direct support costs. From the perspective of an operating command, most support (with the exception of intermediate level maintenance) appears as a free good. Nonetheless, the resource implications of the assumptions that are currently made in the absence of data demand improved requirements estimates. Realistic tests should stress the support system and provide meaningful estimates of both logistics requirements and logistics capability constraints.

This Note does not provide an improved estimate of wartime C-5 engine requirements because of data limitations. It does outline the engineering reasons for expecting engine removal rates, at least for transport aircraft, to be lower in wartime than in peacetime. There may be similar engineering reasons for refuting the linearity assumption for other types of aircraft and equipment.

Undoubtedly the primary reason that the linearity assumption has continued to be used for requirements computation is that it is easy to specify and implement. Requirements statements based on this assumption have survived budget scrutiny for years. Furthermore, the assumption has intuitive appeal--flying should have something to do with demand--and is widely accepted. Changing this assumption, or using assumptions that vary with equipment and application, would make it much more difficult to explain requirements statements.

Current pressures to reduce support resource demands can only intensify as the true costs of supporting new weapons become visible. Much of the attention thus far has focused on the prices being paid for spare and repair parts, but the total cost of spares will remain high even if suppliers can be induced to reduce their prices. Price is only one element of the total cost equation. Other support costs will also continue to be affected by demand forecasts.

Whether or not engineering models are available to predict wartime removal rates, additional effort should be devoted to exploiting potential variability in peacetime aircraft utilization to provide a basis for wartime requirements estimates. Using this natural data source will require behavioral changes within the Air Force. The costs of not making these changes, however, are likely to be reflected in reduced wartime support capability.

This Note has identified a strategy for overcoming the limitations of the available data and improving requirements computations. Buying too much of one item reduces the resources available to procure others. In view of this economic reality, "conservative" assumptions are not really conservative; they merely obscure the true costs of delivering needed wartime capabilities.

Appendix A

SENSITIVITY OF REQUIREMENTS TO PARAMETER ESTIMATES

The sensitivity of wartime support resource requirements to the parameter estimates used in requirements computation is demonstrated by a series of examples. The examples do not consider distributional requirements to simplify exposition.[1] Methods to optimize inventory levels for a multi-site logisitics system are discussed in the literature (e.g., Landi, 1967; Muckstadt, 1973).

Engine requirements are computed to provide 80 percent confidence that an engine will be available when needed, under the assumption that the demand process is Poisson. The total engine requirement consists of "pipeline" stockage, or the expected value of the number of engines tied up in pipelines, and safety stock to protect against variations in demand or resupply. The safety factor to provide an 80 percent confidence level for the Poisson distribution is .8 times the square root of the expected pipeline (Plossl and Wight, 1967). The safety level requirement derived from this relationship must be rounded up to the next integer value to provide at least 80 percent confidence of no stockout.

The pipeline times used in the computation are specified in TO 2-1-18 (Department of the Air Force, 1982). Total times for the base and depot repair cycle are shown in Table A.1.

The "requirement" for spare engines at peak wartime flying levels based on these parameters and a wartime demand rate estimate of .91 removals per flying hours is 171 engines. Distributional considerations, and the fact that some engines will be shipped to overseas locations, account for the difference between this number and the Air Force's computed requirement for 175 engines implied in the text.

The derivation of the 171 "requirement" is shown in Table A.2, which also shows the change in the level of stock required if each of the parameters used in the computation is varied by one unit. The stock

[1]The need to support many different operating locations tends to increase the total stockage requirement.

Table A.1

PIPELINE TIMES USED FOR TF39 ENGINE REQUIREMENTS COMPUTATIONS

Pipeline Segment	Base	Depot
Removal to start work	2	
Base Repair	23[a]	
Removal to shipment		5
Shipping		4[b]
Depot Process/Repair		30
Return shipping		2[b]
Receipt to start work		2
Build up		16
Total	25	67

[a]Peacetime in-work time is 22 days; total wartime repair cycle time is 25 days, assuming 2 shifts, 7 days per week.

[b]Shipping times given are within CONUS. Overseas shipping times are 7 days in each direction. Only the peacetime shipping times are used in the examples to follow. The base return rate (percentage of engines that the basecan repair) is established as 92 percent in the TO.

level is quite sensitive to all of the parameters except the depot repair cycle. This long pipeline has little effect because only 8 percent of the engines flow through it.

Figure A.1 shows the requirement for several different demand rates with all other parameters held constant. These demand rates were developed as follows:

- A: The current demand rate estimate,
- B: Assumes all failures are cycle-driven,
- C: Assumes half of the demand rate is attributable to cycles and half to hour effects,

Table A.2

EFFECTS OF CHANGING PARAMETER VALUES

CASE	DEMAND RATE	DDR	BASE REPAIR TIME	DEPOT REPAIR TIME	NRTS	PIPE LINE DAYS	PIPE LINE QTY	SFTY STOCK	TOT STOCK	CHANGE
Base	.00091	5.69	25	67	.08	28.36	161	10	171	0
Reduced DDR	.0009	5.63	25	67	.08	28.36	160	10	170	-2
Reduced Base Repair Time	.00091	5.69	24	67	.08	27.44	156	10	166	-5
Reduced Depot Repair Time	.00091	5.69	25	66	.08	28.28	161	10	171	0
Reduce NRTS Rate	.00091	5.69	25	67	.07	27.94	159	10	169	-2

DDR: Daily Demand Rate.
NRTS: Not Reparable This Station fraction.

- D: Same assumption as C, but further assumes that the change in cycle density affects only half of the engine "lifetime" because the engines installed at the beginning of a war have expended an average of half their "lives"

Figure A.2 contains requirements estimates based on the same demand assumptions as Fig. A.1 but further assuming a decrease in the wartime base repair cycle to 20 days for the three alternatives. Such a decrease in repair time might be achieved by extending the length of the work day or reducing the depth of repair on some engines. If repair times are distributed exponentially, the average repair time for engines repaired early in a conflict could also be shorter than the average for all engines, particularly if maintenance management chooses to "cherry pick" during the initial stages of the conflict.[2]

[2]"Cherry picking" is a term sometimes used to describe the rational decision to expend repair resources on engines that can be repaired in the shortest amount of time.

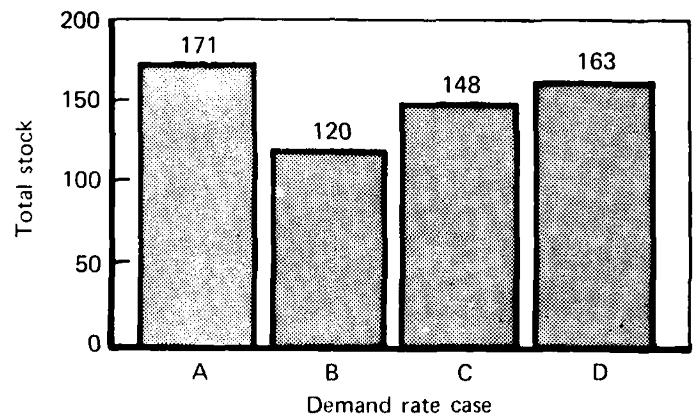


Fig. A.1 — Stockage requirements with varying demand rates

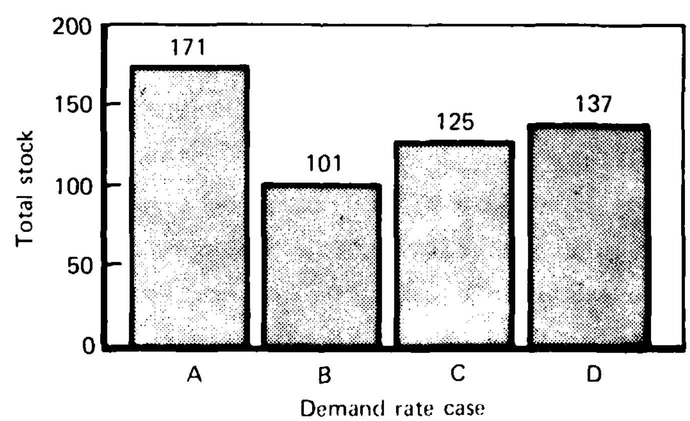


Fig. A.2 — Stockage requirements with varying demand rates and reduced pipeline time

None of these examples take account of the effects of dynamics. Even if demand was proportional to the flying program, the pipeline could not fill until quite late in the scenario. Because the peak wartime flying rate can be sustained for only a short period, the pipeline might never achieve the steady state level implied by this aircraft utilization rate. Considering scenario dynamics explicitly could further reduce spare engine requirements (Berman et al., forthcoming). Improved parameter estimates may even permit the Air Force to avoid further investment in TF39 spares.

The Air Force could take the position that its problem is not one of assessing the validity of the "linearity" assumption, but of determining how many engines are needed to support the expanded C-5 fleet. The results of a complete test are not needed to answer that question.

Current plans are to spend $100 million on additional engines. This "requirement" is based on an estimated 1100 hour MTBR and the pipeline times established in the T.O. (Department of the Air Force, 1982). The demonstration of the sensitivity of requirements statements to parameter estimates provided above reveals that there would be no need for additional engine procurement if the demand rate or pipeline time estimates were reduced by less than 20 percent.

The analysis of survival probabilities presented here and the reliability improvement rationale used to justify the TF39-1C engine modification suggest that the wartime demand rate for this version of the engine should be at least 20 percent lower than the TF39-1A demand rate observed in peacetime. Furthermore, MAC has dedicated transportation for failed engines and can use its maintenance resources more intensively in wartime, so the pipeline times allowed in the T.O. may be excessive. Eliminating the engine buy requirement would release $100 million that could be applied to meet other critical Air Force needs.

Appendix B

THE GE OPSEV MODEL

The manufacturer of the CF6 and TF39 engines has developed an engineering model that has been used successfully to predict commercial engine removal rates in several settings. The basic assumptions embedded in this model are:

- Engine-caused removals are attributable to either functional or potential failure of individual engine components;
- Component failures are due to a combination of cyclic and steady state (flying hour) stress;
- The total failure rate for a component is equal to the sum of its cyclic and steady state failure rates;
- Component failure rates for different mission profiles will differ because cycle density affects the cyclic component of the removal rate;
- The failure rate for the engine as a whole is equal to the sum of the failure rates for these components;
- Increasing the percentage of time that an engine operates at maximum throttle will have a proportional effect on both elements of the failure rate;
- "De-rating" an engine (operating it at less than maximum allowable power) can greatly increase engine life.

The severity ratio developed by the model is simply the failure rate estimated for a particular mission divided by that for a reference mission. The data needed to operate it include:

- The failure rates for the individual components;
- The percentage of the failure rate attributable to cyclic stress;

- The reference and comparison mission profiles (needed to compute cycle density and estimate the percentage of time spent at maximum power);
- The percentage "de-rate" to be used on both missions.

Mission profile descriptions contain estimates of the throttle setting and amount of time spent in each phase of a mission (e.g., taxi, takeoff, cruise, thrust reverse, etc.). These descriptions can be used to produce profile graphs such as those provided in Figs. 3 and 4. The data are used to estimate cycle accrual, which is an important contributor to the removal rate.

The component breakdown, reference mission failure rates, and cyclic fraction of the removal rate used for one model application are shown in Table B.1.

Increasing cycle density increases the cyclic component of the removal rate, and produces a severity ratio greater than unity. Increasing the percentage of total operating time at maximum throttle also increases severity.

As was noted in the text, this model has been used to predict removal rates for turbine engines used in other than aircraft applications. These successful applications of the model in very different environments lend further credence to an engineering approach to predicting removal rates. The results of two such applications are summarized in Table B.2.

The description of the GE OPSEV model provided in this paper is necessarily brief, and has been oversimplified greatly. Further detail is provided in the manufacturer's documentation (Byrd, Barrial, and Kostomay, 1978).

Table B.1

OPSEV MODEL INPUTS

Component/Subassembly	Failure Rate[a]	Cyclic Fraction
Fan Rotor	7.1	58
Fan Stator	12.4	29
LPT Rotor[b]	3.3	86
LPT Stator[b]	4.0	85
Compressor Rotor	4.7	71
Compressor Stator	8.1	46
HPT Rotor[b]	18.3	86
HPT Stator	7.3	72
Combustor	28.2	71
Controls (ambient)	3.6	0
Controls (nonambient)	3.1	0
Controls (cyclic)	2.5	100
Accessories/Gearbox	.1	52
Bearings and Seals (HP)[b]	6.6	0
Bearings and Seals (LP)[b]	15.2	0
Ignition	7.3	100
Fan Frame	5.6	29
Turbine Frame	5.2	94
Inlet Gearbox	.5	52
Total Engine	143.5	55

[a]Failure rates shown in the table are per million flying hours.

[b]"HP" an "LP" refer to High Pressure and Low Pressure. "HPT" stands for "High Pressure Turbine" etc.

Table B.2

OPSEV PREDICTIONS FOR OTHER APPLICATIONS

Application	Predicted Rate	Observed Rate
Ship propulsion	.268	.286
Gas pumper	.114	.09

Appendix C

DATA PROCESSING

This appendix provides additional detail concerning the procedures used to process the data used in the exploratory and statistical analyses as well as several observations concerning the data collected for the analysis.

DESCRIPTION OF THE DATA

As was indicated in Sec. IV, extracts from two files maintained in the GPS system were used to create the working datasets:

- Aircraft flying histories,
- Engine installation and removal histories.

The aircraft utilization file records several utilization measures, including:

- Flying hours,
- Sorties,
- Full-stop landings,
- Total landings,
- Total landing gear cycles.

These parameters are recorded by (1) aircraft tail number, (2) location, and (3) day. Monthly summary records contain the same utilization parameters summed across operating locations, as well as monthly and cumulative totals for each of the utilization measures. These data are also maintained in the Aerospace Vehicle Status and Utilization Reporting System (Department of the Air Force, 1976).

The Engine Status Reporting System maintained at the Oklahoma City Air Logistics Center is updated by both periodic inventory reconciliation and status change reports (Department of the Air Force, 1975). In addition to the quarterly reports of engine status and hours, this system is updated whenever an engine is installed, removed, or transferred to another accountable organization. The latter transactions are also recorded in the GPS system.

Both installation and removal reports record:

- Aircraft tail number,
- Engine position,
- The transaction date,
- Engine hours at the time of the transaction.

Removal records also contain a code indicating the reason for removal, which describe the symptoms that led to a decision to remove the engine rather than the cause of failure. Some of these removal reasons are not maintenance-significant. For example, an engine may be removed to facilitate other aircraft maintenance or to fill a hole in another aircraft. However, the aircraft, transaction date, and engine hour data can be used to link engine and aircraft records so that it is possible to include the more complete aircraft utilization data in engine lifetime histories.

The flying history records obtained for this study cover the period from late 1972 through the end of 1982. Engine removal records were available from 1969 through 1982. Not all of these data could be used, however, because several pre-1976 engine histories were incomplete.

DATA PROCESSING

The engine utilization and engine history records were merged by aircraft tail number to create engine lifetime records, which contain all of the aircraft utilization parameters described above for each engine installed since 1976.

Although most data deficiencies were eliminated by excluding the pre-1976 data, some problems remained. There were a few inconsistencies in reported engine installation and removal dates. In addition, some of the flying records contained obvious errors. For example, some records recorded more sorties than full-stop landings.

The engine history problems resulted from an effort to improve GPS data quality several years ago. Although not all transaction dates were corrected, engine hours were recorded accurately. Thus engine hours were used in conjunction with the flying records to determine the true installation date. Because every aircraft has only four engines installed when it operates, reconstructing these engine installation histories was tedious but straightforward.

Air Force personnel familiar with the GPS system indicated that some of the problems encountered in the flying records were attributable to the fact that landing data were not always recorded accurately in the mid-1970s. The sortie data were considered to be more reliable, so the data were edited to ensure that the number of landings was at least equal to the number of reported sorties. Fewer than 5 percent of the engine utilization records were affected by this edit. Its effect, if any, was to increase the average cycle length computed for the affected engines in the analyses described in Secs. VI and VII.

The final step in this preliminary processing was to sum the relevant utilization parameters across engine records when the removal reason code indicated that the removal was not maintenance-significant. Thus the records contain utilization data from installation until the engine was removed for cause. "Censored" records--for engines that had not been removed for cause by the end of 1982--contain the hours, sorties, etc. accrued on the engine since installation.

These same data were used to create the detailed utilization records for each month an engine was installed. The removal records were used as the key to extract monthly summary records for the engine's lifetime. Flying hours were allocated for the month of installation and removal on the assumption that monthly flying hours were spread equally across the days in the month.

DATA OBSERVATIONS

The TF39-1A engine has had a maximum Time Between Overhauls (TBO) of 5000 hours for several years. The TBO limit, which is a policy variable, could censor some of the data if time-change limits resulted in engine removal before performance deterioration or failure required the engine to be removed.

There were no "maximum time" removals in the data. In fact, there were several examples of engines that never "aged" more than 2000 hours because they were overhauled repeatedly. This suggests either that some engines are "lemons" or that engines are being evacuated to the depot when they could be repaired by JEIM. Some recent data undoubtedly included records for engines forced to the depot for conversion to the TF39-1C configuration, but this factor alone cannot explain all of these apparently premature overhauls.

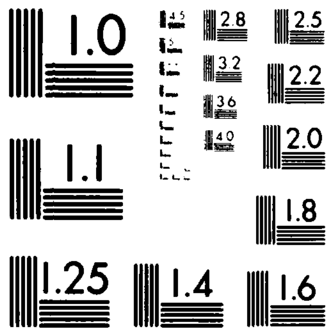
Some other engines were removed for cause several times after accumulating only a few hours. These could also have been lemons--or else maintenance repeatedly failed to correct the problem that prompted the initial removal.

Despite these problems, compared with the data contained in other Air Force data systems, GPS data quality was surprisingly good. The errors that did exist were corrected quite easily. One factor that differentiates GPS from most other data systems is that it is used for base maintenance management as well as to collect maintenance data. The user is motivated to provide accurate inputs by the fact that the outputs affect his ability to perform his job.

One of the common problems with centralized data collection systems is that the people responsible for updating the system do not realize any benefit for their efforts. Those responsible for the GPS system design appear to have addressed this problem by making the reporting system a local management system. This helps to account for the relatively high quality of the data.

AD-A147 923 ESTIMATING WARTIME SUPPORT RESOURCE REQUIREMENTS STATISTICAL AND RELATED POLICY ISSUES(U) RAND CORP SANTA MONICA CA L B EMBRY JUL 84 RAND/N-2163-AF 2/2

UNCLASSIFIED F49620-82-C-0018 F/G 15/5 NL



MICROCOPY RESOLUTION TEST CHART

Appendix D

BAYESIAN ANALYSIS OF REMOVAL DATA

Bayesian methods can be used to update estimates of the engine removal rate during the course of an experiment designed to test the effect of cycle length on the removal rate. The example provided below follows the approach previously documented in the reliability literature (Soland, 1968, 1969).

Assume that the GE model yields a wartime removal rate estimate for the TF39-1C of .0007 (before adjusting for increased wartime time at maximum power). This corresponds to an MTBR of about 1430 hours.

Assume further that removals are believed to be generated by a Weibull process with parameters α and $\zeta = \eta^{-\alpha}$, $\gamma = 0$. Parameter values are unknown. Uncertainty concerning their values will be expressed by assuming a discrete distribution for α and a continuous distribution for ζ conditional on the distribution of α.

The parameter values are not of direct interest but will be useful for deriving a posterior mean life estimate for the Weibull distribution (a prior estimate was obtained with the model). Although product improvements tend to reduce the slope of the hazard function for turbine engines (Nowlan and Heap, 1974), the 1C modification is unlikely to result in an exponential distribution of TF 39-1C life. Bayesian procedures will be used to update estimates of the Weibull parameters and through them the mean life estimate.

The shape parameter is assumed to lie in the range $1.2 \leq \alpha \leq 1.3$ based on past data and the expected results of the engine modification program. Probabilities are assigned to discrete values as follows:[1]

α_i	$Pr(\alpha_i = \alpha)$
1.2	.4
1.3	.6

[1]An infinite number of values of α could be assumed. This example uses only two to simplify exposition.

It is expected that median life will be at least 1100 hours,[2] and that there is a 25 percent chance that an engine will last more than 2000 hours.

Two aircraft (eight engines) are put on test at the beginning of the first month, and two more are added at the end of the month. Test engines accrue 150 hours per month. By the end of the fourth month, three of the 16 engines have been removed at $t_k = 90, 280$, and 550. The first and last removals occurred on the initial two aircraft.

The scale parameter is assumed to be known conditional on α. A "kernel" of likelihood (Raiffa and Schlaifer, 1961) is given by:

$$\zeta^r e^{-\zeta y}$$

$$\text{where } y = \sum_{j=1}^{n} t_j^{\alpha}$$

This corresponds to a gamma -1 density:

$$f_{\gamma 1}(\zeta | r, y) = \frac{y^r \zeta^{r-1} e^{-\zeta y}}{\Gamma(r)}, \quad \begin{array}{l} 0 \le \zeta \le \alpha \\ 0 \le r,\ y \le \alpha \end{array}$$

where r is the number of failures observed, or the imputed number of failures in a prior analysis. This parameterization has the interesting property that the "assumed" experience implicit in a prior estimte is made explicit through the values of the parameters r and y.

Prior estimates of r and y can be computed numerically using the expression:

---

[2]Median life for an exponential distribution would be about 1000 hours.

$$\sum_{\zeta|\alpha_i}' \left[1 - F_w(t|\zeta,\alpha)\right] = \left[y_i' / (y_i' + t^{\alpha_i})\right]^{r_i'}$$

(The notation E', y', r' refers to the prior estimates).

These values for the data available before the experiment are:

α_i	r_i	y_i
1.2	14.46	90891
1.3	4.29	51276

Soland defines an additional statistic $v = \prod_{k=1}^{r} t_k$

where t_k are the times of the r failures. The likelihood of the evidence given the preposterior estimates of the Weibull parameters is then given by:

$$\ell(z|\zeta,\alpha_i) = \alpha_i^r \zeta^r v^{\alpha_i - 1} e^{-\zeta y_i}$$

and a posterior distribution of α can be developed using Bayes' theorem as follows:

$$\Pr\left\{\tilde{\alpha} = \alpha_i | \zeta\right\} = p_i'' = \frac{p_i' \alpha_i^r v^{\alpha_i} y_i'^{r_i'} \Gamma\left(r_i''\right) \Big/ \left[y_i''^{r_i''} \Gamma\left(r_i'\right)\right]}{\sum_{i=1}^{m} p_i' \alpha_i^r v^{\alpha_i} y_i'^{r_i'} \Gamma\left(r_i''\right) \Big/ \left[y_i''^{r_i''} \Gamma\left(r_i'\right)\right]}$$

where $r_i'' = r_i' + r$, $y_i'' = y_i' + y_i$.

The posterior distribution of ζ conditional on α_i is:

$$f\left(\zeta \mid \alpha_i\right) = f_{Y_i}\left(\zeta \mid r_i'', y_i''\right) = \frac{y_i''^{\,r_i''} \zeta^{r_i''-1} e^{-\zeta y_i''}}{\Gamma\left(r_i''\right)}$$

(The superscript " refers to the posterior distributions).

The data accumulated from the early stages of the experiment described above yields the following statistics:

$$\begin{aligned} r &= 3 \\ y_1 &= 24459^{3} \\ y_2 &= 45195 \\ v &= (90)(280)(550) = 1.386 \times 10^{7} \end{aligned}$$

These statistics yield a posterior distribution on α. The values are tabulated below:

i	α	$Pr(\alpha_i = \alpha)$
1	1.2	.39
2	1.3	.61

A posterior estimate of mean life can then be obtained based on the conditional gamma -1 density assumed for ζ, weighted by the appropriate values of α_i.

[3]
$$\begin{aligned} y_i = {} & (90)^{\alpha_i} + (280)^{\alpha_i} + (550)^{\alpha_i} && \text{(failures)} \\ & + (510)^{\alpha_i} + (20)^{\alpha_i} + (50)^{\alpha_i} && \text{(replacements)} \\ & + 6(600)^{\alpha_i} + 7(300)^{\alpha_i} && \text{(non-failures)} \end{aligned}$$

For each α_i, the expression for expected mean life (ML) is:

$$ML_i = \frac{\Gamma(1 + 1/\alpha_i)(y_i'')^{1/\alpha_i}\,\Gamma(r_i'' - 1/\alpha_i)}{(r_i'')}$$

The resulting estimates are:

$ML_1 = 1525$

$ML_2 = 1280$

$ML = 1375$

These same procedures could be used to combine the predictions of the GE model with operational experience in a real experiment.

BIBLIOGRAPHY

Altman, O.L. and C.G. Goor, "Actuarial Analysis of the Operating Life of B-29 Aircraft Engines," *Journal of the American Statistical Association, 48*, 1948, pp. 190-203.

American Airlines, Inc., *The VALUE of Maintenance*, presented at the FAA Maintenance Symposium, Oklahoma City, November 1967.

Armitage, P., "The Role of Randomization in Clinical Trials," *Statistics in Medicine, 1*, 1982, pp. 345-352.

Barlow, R.E. and R.A. Campo, "Total Time on Test Processes and Applications to Failure Data Analyses," *Reliability and Fault Tree Analysis: Theoretical and Applied Aspects of System Reliability and Safety Assessment*, SIAM, Philadelphia, 1975, pp. 451-481.

Barlow, R.E. and F. Proschan, *Statistical Theory of Reliability and Life Testing: Probability Models*, Holt, Rinehart and Winston, Inc., New York, 1975.

Barlow, R.E. and E.M. Schener, "Estimation from Accelerated Life Tests," *Technometrics, 13-1*, February 1971, pp. 145-159.

Berman, M.B., T.F. Lippiatt, E. Poggio, and R. Sims, *Predicting Aircraft Component Removal Rates with Varying Flying Programs: A TF33-7 Engine Case Study*, The Rand Corporation, unpublished research.

Berrettoni, J.N., "Practical Applications of the Weibull Distribution," ASQC Convention Transactions, Cincinnatti, 1962, pp. 303-323.

Birkler, J. L. and J. R. Nelson, "Aircraft Turbine Engine Monitoring Systems: Overview and Lessons Learned from Six Case Studies," The Rand Corporation, P-6337, May 1979.

Breslow, N., "Statistical Methods for Censored Survival Data," *Environmental Health Perspectives, 32*, 1979, pp. 181-192.

Byrd, F., H. Barrial, and D. Kostomay, *Turbine Engine Life Development Definition: Final Report*, General Electric Aircraft Engine Group, AFAPL-TR-78-45, June 1978.

Carlyle, L. *Comparison of Military and Commercial Design-to-Cost Aircraft Procurement and Operational Support Practices*, McDonnell . Douglas Corp., Long Beach, AFFDL-TR-75-147, January 1976.

Casey, D.L., *Logistics Impact of Longer C-5 Missions*, AFLC Technical Report No. 76-16, May 1977.

Chorofas, D.N., *Statistical Processes and Reliability Engineering*, D. Van Nostrand Co., Inc., Princeton, 1960.

Coppola, A., *"Bayesian" Reliability Tests Made Practical*, Rome Air Development Center, RADC-TR-81-106, July 1981.

Cote, S. M. and J. L. Birkler, *Simulated Mission Endurance Test (SMET) for an Aircraft Engine to be Used in a Fighter/Attack Role*, Naval Air Development Center, NADC-77031-30, April 1979.

Davis, N.K., *United Airlines Test and Replace as Necessary (TARAN) Program*, presented at the FAA Maintenance Symposium, Oklahoma City, OK, November 1967.

Department of the Air Force, *Aerospace Vehicle Status and Utilization Reporting System*, AFR-65-110, 1976.

Department of the Air Force, *Aircraft Engine and Module Management by Operating Limits and Pipeline Times*, TO 2-1-18, July 1982.

Department of the Air Force, "An Analysis of the Aircraft Replenishment Spares Acquisition Process: Report of the Corona Require Study Group," March 1983.

Department of the Air Force, *C-5 Ground Processing System*, TO 1-1G-1, 1981.

Department of the Air Force, *Engine Status Reporting*, AFM-400-1, August 1975.

Department of Defense, *Reliability Design Qualification and Production Acceptance Tests: Exponential Distribution*, MIL-STD-781C, October 1977.

Donaldson, T.S. and A.F.M. Sweetland, *The Relationship of Flight-Line Maintenance Man Hours to Aircraft Flying Hours*, The Rand Corporation, RM-5701-PR, August 1968.

Drezner, S.M., "The Analogs 80 Project: Reallocation of Activities Alternative," The Rand Corporation, unpublished briefing, 1975.

Embry, L.B., "Estimating Wartime Depot Workload," The Rand Corporation, unpublished briefing based on results of analysis of the landing gear facility at Ogden ALC, 1982.

Embry, L.B. and G.B. Crawford, "Extrapolating Peacetime Demand Experience to Wartime Activity Levels: The 'Linearity' Issue," Published in the Proceedings of the USAF Capability Assessment Symposium (LOGCAS 83), Air Force Logistics Management Center, AFLMC Report 781029-4, May 1983.

Feeney, G. J., J. W. Petersen and C. C. Sherbrooke, "An Aggregate Base Stockage Policy for Recoverable Spare Parts," The Rand Corporation, RM-3644-PR, June 1963.

Freedman, L.S., "Tables of the Number of Patients Required in Clinical Trials using the Logrank Test," *Statistics in Medicine, 1*, 1982, pp. 121-129.

Gandara, A. and M.D. Rich, *Reliability Improvement Warranties for Military Procurement*, The Rand Corporation, R-2264-AF, December 1977.

Gans, D.J., "A Simple Method Based on Broken-line Interpolation for Displaying Data from Long-Term Clinical Trials," *Statistics in Medicine, 1*, 1982, pp. 131-137.

Groover, C.W., "Resource Allocation for Readiness and Sustainability," unpublished briefing delivered at the Industrial College of the Armed Forces, 1984.

Hahn, G.J. and S.S. Shapiro, *Statstical Models in Engineering*, John Wiley & Sons, Inc., New York, 1967.

Hall, W.J. and J.A. Wellner, "Confidence Bands for a Survival Curve from Censored Data," *Biometrika, 67-1*, 1980, pp. 133-143.

Hirschman, A.O., *Exit, Voice and Loyalty*, Harvard University Press, Cambridge, 1970.

Hitch, C.J. and R.N. McKean, *The Economics of Defense in the Nuclear Age*, The Rand Corporation, R-346, March 1960.

Hunsaker, H.D. et al., *Logistics Impact of Sortie Oriented Aircrew Training*, AFLC Technical Report No. 76-7, April 1977.

Jones, T.V., "Logistics and the Military End Game," *Defense Management Journal, 19*, December 1983, pp. 12-16.

Kahn, H.D., "Least Squares Analysis of Accelerated Life Tests for the Power Rule and Arrhenius Models," *The Theory and Applications of Reliability with Emphasis on Bayesian and Nonparametric Methods*, Volume II, Academic Press, New York, 1977, pp. 381-400.

Kalbfleisch, J.D. and R.L. Prentice, *The Statistical Analysis of Failure Time Data*, John Wiley and Sons, New York, 1980.

Kao, J.K., "A Graphical Estimation of Mixed Weibull Parameters in Life-Testing of Electron Tubes," *Technometrics, 1-4*, November 1959, pp. 389-407.

Kaplan, E.L. and P. Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association, 53*, June 1958, pp. 457-481.

Kay, R. and M. Schumacher, Unbiased Assessment of Treatment Effects on Disease Recurrence and Survival in Clinical Trials, *Statistics in Medicine*, *2*, 1983, pp. 41-58.

Kern, G.A., and T.M. Darns, *Operational Influences on Reliability*, Rome Air Development Center, AFSC, RADC-TR-76-366, December 1966.

Landi, D.M., *Positioning Recoverable Spares in Military Airlift Networks*, The Rand Corporation, RM-5087-PR, March 1967.

Lippiatt, T.F., "Analysis of Maintenance Data from the SALTY ROOSTER Exercise," unpublished briefing, The Rand Corporation, 1978.

Mann, N.R., *Extension of Results Concerning Parameter Estimators, Tolerance Bounds and Warranty Periods for Weibull Models*, Rocketdyne, Canoga Park, Ca., ARL 70-0010, January 1970.

Miller, D.R. and N.D. Singpurwalla, *Failure Rate Estimation Using Random Smoothing*, George Washington University School of Engineering and Applied Science, Washington, D.C., T-347, April 1977.

Muckstadt, J.A., "A Model for a Multi-Item Multi-Echelon Inventory System," *Management Science*, *20-4*, December 1973, pp. 472-481.

Mullins, General J.P., "Establishing Supportability as a Critical Requirements Factor," *Defense Management Journal*, *19-4*, December 1983, pp. 2-6.

National Aeronautics and Space Administration (NASA) Lewis Research Center, *Aircraft Engine Diagnostics*, Proceedings of a Conference held in Cleveland, Ohio, May 1981.

Nelson, L.S., "Weibull Probability Paper," *Evaluation Engineering*, May/June 1967.

Nelson, Wayne, "Hazard Plotting for Incomplete Failure Data," *Journal of Quality Technology*, *1-1*, January 1969, pp. 27-52.

Nowlan, F.S. and H. Heap, *Reliability-Centered Maintenance*, Office of the Assistant Secretary of Defense (Manpower, Reserve Affairs and Logistics), Washington, D.C., 1974.

Parzen, E., "On Estimation of a Probability Density Function and Mode," *Ann. Math. Statist.*, *33*, 1962, pp. 1056-1076.

Peto, R., M.C. Pike, P. Armitage, N.E. Breslow, D.R. Cox, S.V. Howard, N. Mantel, K. McPherson, J. Peto, and P.G. Smith, "Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient I. Introduction and Design," *British Journal of Cancer*, *34*, 1976, pp. 585-612.

Peto, R., M.C. Pike, P. Armitage, N.E. Breslow, D.R. Cox, S.V. Howard, N. Mantel, K. McPherson, J. Peto, and P.G. Smith, "Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient II. Analysis and Examples," *British Journal of Cancer, 35*, 1975, pp. 1-39.

Plossl, G.W. and O.W. Wight, *Production and Inventory Control*, Prentice-Hall, Inc., Englewood Cliffs, N.J., 1967.

Raiffa, H., *Decision Analysis*, Addison-Wesley Publishing Company, Reading, Mass., 1968.

Raiffa, H. and R. Schlaifer, *Applied Statistical Decision Theory*, Harvard Graduate School of Business Administration, Boston, 1961.

Rice, D.B., *Defense Resource Management Study: Final Report*, Washington, D.C., 1979.

Rice, J. and M. Rosenblatt, "Estimation of the Log Survivor Function and Hazard Function," *The Indian Journal of Statistics, 38-A-1*, 1976, pp. 60-78.

Rich, M.D. and S.M. Drezner, "An Integrated View on Improving Combat Readiness," The Rand Corporation, N-1797-AF, February 1982.

Rogers, W.H. and J. Hanley, "Weibull Regression and Hazard Estimation," *SAS Users Group International Conference Proceedings*, February 1982.

Sallee, G. Philip, *Economic Effects of Propulsion System Technology on Existing and Future Transport Aircraft*, American Airlines Inc., NASA CR 134645, July 1974.

Shaw, Col. C.C., *The Effect of System Operation Duration on Failure Rate: A Study of Strategic Airlift Aircraft*, U.S. Air Force ACOS Studies & Analysis, December 1982.

Shimi, I. and C.P. Tsokos, "The Bayesian and Nonparametric Approach to Reliability Studies: A Survey of Recent Work," *The Theory and Applications of Reliability with Emphasis on Bayesian and Nonparametric Methods*, Volume II, Academic Press, New York, 1977, pp. 5-47.

Shulman, H. and J. Gebman, "Maturational Development," unpublished briefing, The Rand Corporation, 1981.

Shulman, H., J. Gebman and N.Y. Moore, "F-15 AIS Study," unpublished briefing, The Rand Corporation, 1980.

Shurman, Maurice, *B-52D Operations - Southeast Asia vs. CONUS*, Boeing Military Airplane Systems Division, D162-10015-1, September 1970.

Shurman, Maurice, "Time-Dependent Failure Rates for Aircraft Components," unpublished briefing, Boeing Aerospace Company, 1979.

Siegel, S., *Nonparametric Statistics for the Behavioral Sciences*, McGraw-Hill Book Co., Inc., New York, 1956.

Simon, R. and Y.J. Lee, "Nonparametric Confidence Limits for Survival Probabilities and Median Survival Time," *Cancer Treatment Reports, 66-1*, January 1982, pp. 37-42.

Singpurwalla, N.D. and F.A. Al-Khayyal, "Accelerated Life Tests Using the Power Law Model for the Weibull Distribution," *The Theory and Applications of Reliability with Emphasis on Bayesian and Nonparametric Methods*, Volume II, Academic Press, New York, 1977, pp. 381-400.

Soland, R.M., "Bayesian Analysis of the Weibull Process with Unknown Scale Parameter and Its Application to Acceptance Sampling," *IEEE Transactions on Reliability, R-17-2*, June 1968, pp. 84-90.

Soland, R.M., "Bayesian Analysis of the Weibull Process with Unknown Scale and Shape Parameters," *IEEE Transactions on Reliability, R-18-4*, November 1969, pp. 181-184.

Tribus, Myron, *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York, 1967.

Tukey, J.W., *Exploratory Data Analysis*, Addison-Wesley Publishing Co., Reading, Mass., 1977.

Walker, G.I., *Logistics Forecasting for Achieving Low Life Cycle Costs*, Advisory Group for Aerospace Research and Development, AGARD-LS-107, May 1980.

Walker, W.E., J.M. Chaiken and E.J. Ignall (eds.), *Fire Department Deployment Analysis*, North Holland, New York, 1979.

Weibull, W., "A Statistical Distribution Function of Wide Applicability," *Journal of Applied Mechanics, 18*, 1951, pp. 293-297.

Wulf, R. N., *Engine Diagnostics Program: CF6-50 Engine Performance Deterioration*, General Electric Company Aircraft Engine Business Group, NASA CR-159867, November 1980.